Variational Methods for Evolution

DISSERTATION

zur Erlangung des akademischen Grades

Dr. rer. nat.
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Dipl.-Math. Matthias Liero

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Alexander Mielke
2. Prof. Dr. Ulisse Stefanelli
3. Priv.-Doz. Dr. Annegret Glitzky

eingereicht am: 20. Juli 2012
Tag der mündlichen Prüfung: 07. Dezember 2012

Abstract

This thesis deals with the application of variational methods to evolution problems governed by partial differential equations. In particular, the first part of this work is devoted to systems of reaction-diffusion equations that can be formulated as gradient systems with respect to an entropy functional and a dissipation metric. The dissipation metric is given in terms of a so-called Onsager operator, which is a sum of a diffusion part of Wasserstein type and a reaction part. After giving a brief survey of the framework for gradient systems developed by A. Mielke we provide methods for establishing geodesic λ-convexity of the entropy functional by purely differential methods. Thus we circumvent arguments from mass transportation, which are not available for systems of equations or even for scalar equations with reaction terms since mass is in general not conserved. Geodesic λ-convexity is beneficial, however, it is a strong structural property of a gradient system that is rather difficult to achieve. Several examples, including a drift-diffusion system, provide a survey on the applicability of the theory.

Next, we demonstrate the application of variational methods, such as Mosco and Γ-convergence, to derive effective limit models for multiscale problems. The crucial point in this investigation is that we rely only on the gradient structure of the systems. We consider two model problems: The rigorous derivation of an Allen-Cahn system with bulk/surface coupling and of an interface condition for a one-dimensional diffusion equation. The derivation of these limit systems is based on the energy-dissipation equation or the De Giorgi principle for gradient flows, which characterizes solutions as curves of maximal slope.

The second part of this thesis is devoted to the so-called Weighted-Inertia-Dissipation-Energy principle (abbreviated WIDE principle in the following). The WIDE principle is a global-in-time variational principle for evolution equations either of conservative or dissipative type. It relies on the minimization of a specific parameter-dependent family of functionals (WIDE functionals) with minimizers characterizing entire trajectories of the system. We prove that minimizers of the WIDE functional converge, up to subsequences, to weak solutions of the limiting PDE when the parameter tends to zero. Thus, this variational principle may serve as a selection criterion in case of nonuniqueness of solutions.

Here we distinguish between two cases: the finite and infinite time horizon case. Each case is treated by two completely different approaches: For the finite time horizon case we use a quite technical time-discretization scheme, which is of interest in its own, while in the second case we use a rescaling and reparametrizations of time to deduce the crucial a priori bounds on the minimizers. The latter then allows us to select converging subsequences and to pass to the limit in the Euler-Lagrange equations for the WIDE functionals.

The interest for this perspective is that of moving the successful machinery of the Calculus of Variations (Direct Method, Γ-convergence, relaxation) to evolutionary situations. Moreover, we are able to treat dissipative and nondissipative situations simultaneously. Notably, the WIDE principle allows for a rigorous connection between these two regimes by means of Γ-convergence.

In the case of semilinear wave equations the WIDE principle corresponds to a long-standing conjecture by E. De Giorgi, which was just recently proved.

Zusammenfassung

Das Thema dieser Dissertation ist die Anwendung von Variationsmethoden auf Evolutionsgleichungen parabolischen und hyperbolischen Typs. Im ersten Teil der Arbeit beschäftigen wir uns mit Reaktions-Diffusions-Systemen, die sich als Gradientensysteme schreiben lassen. Hierbei verstehen wir unter einem Gradientensystem ein Tripel bestehend aus einem Zustandsraum, einem Entropiefunktional und einer Dissipationsmetrik. Die Dissipationsmetrik ist durch einen Onsager-Operator gegeben und setzt sich aus einem Diffusions- und einem Reaktionsteil zusammen, wobei ersterer vom Wasserstein-Typ ist. Nach einer kurzen Zusammenfassung des von A. Mielke aufgestellten Formalismus für Gradientensysteme geben wir Bedingungen an, die die geodätische λ-Konvexität des Entropiefunktionals sichern. Diese Bedingungen sind abstrakt bezüglich des Onsager-Operators formuliert. Insbesondere benutzen wir hier nicht das Prinzip des Optimalen Transports, das aufgrund fehlender Massenerhaltung bei Systemen mit Reaktionstermen nicht anwendbar ist. Geodätische λ-Konvexität ist eine wertvolle aber auch starke strukturelle Eigenschaft und relativ schwer zu zeigen. Wir zeigen anhand zahlreicher Beispiele, darunter ein Drift-Diffusions-System, dass dennoch interessante Systeme existieren, die diese Eigenschaft besitzen.

Einen weiteren Punkt dieser Arbeit stellt die Anwendung von Konvergenzbegriffen wie Mosco- und Γ-Konvergenz auf Gradientensysteme dar. Wir betrachten hierbei zwei Modellsysteme aus dem Bereich der Mehrskalenprobleme: Erstens, die rigorose Herleitung einer Allen-Cahn-Gleichung mit dynamischen Randbedingungen und zweitens, einer Interface-Bedingung für eine eindimensionale Diffusionsgleichung jeweils aus einem reinen Bulk-System. Wir benutzen hierbei das De Giorgi-Prinzip für Gradientensysteme, das Trajektorien des Systems als Kurven maximaler Steigung (curves of maximal slope) charakterisiert und in der Form einer Energie-Dissipations-Gleichung geschrieben ist.

Im zweiten Teil der Arbeit beschäftigen wir uns mit dem sog. Weighted-Inertia-Dissipation-Energy-Prinzip (WIDE-Prinzip) für Evolutionsgleichungen. Hierbei werden Trajektorien eines Systems als (Grenzwerte von) Minimierer(n) einer parametrisierten Familie von Funktionalen charakterisiert. Dies erlaubt es, Werkzeuge aus der Theorie der Variationsrechung (Direkte Methode, Γ-Konvergenz, usw.) auf Evolutionsprobleme anzuwenden. Die Euler-Lagrange-Gleichungen dieser WIDE-Funktionale können als elliptische Regularisierung der Ausgangsgleichung interpretiert werden. Wir zeigen, dass Minimierer (bzw. stationäre Punkte) der WIDE-Funktionale gegen Lösungen des Ausgangsproblems konvergieren. Hierbei betrachten wir getrennt voneinander den Fall des beschränkten und des unbeschränkten Zeitintervalls, die jeweils mit verschiedenen Methoden behandelt werden. Während wir im ersten Fall ein zeitdiskretes Hilfsproblem untersuchen, benutzen wir im zweiten Fall Zeitreparametrisierungen, um gleichmäßige Schranken für die Minimierer der WIDE-Funktionale herzuleiten. Diese Schranken erlauben die Auswahl einer konvergenten Teilfolge, sodass wir in der Euler-Lagrange-Gleichung für die WIDE-Funktionale zum Grenzwert übergehen können. Insbesondere stellt das WIDE-Prinzip ein Auswahlkriterium dar, wenn keine Eindeutigkeit der Lösungen des Ausgangsproblems vorliegt. Ferner erlaubt uns das WIDE-Prinzip dissipative und nicht-dissipative Systeme zu betrachten und ihre Beziehung mit Hilfe der Γ-Konvergenz zu untersuchen. Im Fall der semilinearen Wellengleichung entspricht das WIDE-Prinzip einer Vermutung von E. De Giorgi, die erst vor kurzem bewiesen wurde.

Acknowledgment

First and foremost, I would like to express deep gratitude to Dr. Annegret Glitzky, Professor Alexander Mielke and Professor Ulisse Stefanelli. They always supported and encouraged me in my work and provided constructive criticism and invaluable guidance.

Moreover, I am very grateful to the many people with whom I have had the pleasure to discuss not only mathematical ideas – notably the people of Research group Partial Differential Equations at Weierstrass Institute. Also, I would like to thank the people at IMATI Pavia for being helpful and for the warm hospitality during my stay.

The work leading to the results in this thesis has received funding from Deutsche Forschungsgemeinschaft within the Research Training Group 1128, Analysis, Numerics, and Optimization of Multiphase Problems, and the MATHEON Research Center under Project D22. Moreover, during the stay at IMATI Pavia I received funding by the ERC programme Mathematics for Shape Memory Technologies in Biomechanics (FP7-IDEAS-ERC-StG Grant #200947) which I gratefully acknowledge.

Contents

Part I

Gradient systems

1 Introduction to Part I

As a paradigm of dissipative evolution, the class of gradient flows has attracted constant attention during the last decades (see [Kom67, CrP69, Bré71, Bré73]). A major step was taken in the late nineties with the introduction of *Wasserstein gradient flows* by JORDAN, KINDERLEHRER, and OTTO [JKO98, Ott98, Ott01]. Since then it has become clear that a large number of diffusion equations can be written as (metric) Wasserstein gradient flows (see e.g. [AGS05]). Recently, MIELKE [Mie11b] succeeded in carrying the basic ideas of the Wasserstein setting over to reaction-diffusion and reaction-drift-diffusion systems and therefore opening them for the application of variational methods.

Simply put, gradient flows are evolutionary systems driven by an energy, in the sense that the energy decreases along solutions as fast as possible. In order to specify what "as fast as possible" means, one defines a dissipation mechanism that characterizes the decrease of energy along solutions. Here, we adopt the framework for *gradient systems* established in [Mie11b] (see also [GlM13, Mie13]), where a gradient system is understood as a triple $(X, \mathcal{E}, \mathcal{G})$. The *state space* X contains the states $u \in X$, $\mathcal{E} : X \to \mathbb{R}$ is the (differentiable) *driving functional* and $\mathcal{G}(u) : X \to X^*$ is a linear, symmetric and positive (semi-)definite *metric tensor* which represents the dissipative structure of the system.

The gradient flow associated with the gradient system $(X, \mathcal{E}, \mathcal{G})$ is then given as the abstract force balance

$$\mathcal{G}(u)\dot{u} = -\mathrm{D}\mathcal{E}(u) \qquad \Longleftrightarrow \qquad \dot{u} = -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u), \tag{1}$$

where $\mathcal{K}(u) \stackrel{\text{def}}{=} \mathcal{G}(u)^{-1}$ denotes the inverse operator. We call $\mathcal{K}$ *Onsager operator*, being also symmetric and positive (semi-)definite, and the triple $(X, \mathcal{E}, \mathcal{K})$ *Onsager system*. Here, the naming refers to ONSAGER's fundamental symmetry relations, meaning $\mathcal{K} = \mathcal{K}^*$, and the Onsager principle (see [Ons31]).

We will learn in Chapter 2 that from the modeling perspective it is convenient to consider the Onsager operator $\mathcal{K}$ instead of $\mathcal{G}$: Often differential equations are written in rate form $\dot{u} = \mathcal{F}(u)$, where the vector field $\mathcal{F}$ is additively decomposed into different physical phenomena (e.g. diffusion, reaction). This additive split can also be used for the Onsager operator, as long as all the different effects are driven by the same functional.

The class of evolution systems that can be written as a gradient system $(X, \mathcal{E}, \mathcal{G})$ is rich: e.g. general reaction-diffusion systems, with reactions satisfying the *detailed balance condition*, can be written as a gradient system with respect to the relative entropy (see subsequent sections). Moreover, in [GlM13] the application to systems with bulk/interface coupling was shown, which is of great interest e.g. in the theory of heterostructure semiconductor devices.

Throughout the following chapters we are mainly interested in reaction-diffusion systems which can be written as Onsager system $(X, \mathcal{E}, \mathcal{K})$, where the latter is of the form

$$\begin{aligned} &\mathcal{E}(\boldsymbol{u}) = \textstyle\int_\Omega E(x, u(x))\,\mathrm{d}x, \qquad \mathcal{K}(u) = \mathcal{K}_{\text{diff}}(u) + \mathcal{K}_{\text{react}}(u), \\ &\text{where} \quad \mathcal{K}_{\text{diff}}(u)\xi = -\mathrm{div}(\mathbb{M}(u)\nabla\xi), \quad \text{and} \quad \mathcal{K}_{\text{react}}(u)\xi = \mathbb{K}(u)\xi. \end{aligned}$$

Here, $\mathbb{M}(x,u)$ and $\mathbb{K}(x,u)$ are symmetric, positive semidefinite tensors of order four and two, respectively. The corresponding evolution equation (1) then reads

$$\dot{u} = -\mathrm{div}\Big(\mathbb{M}(x,u)\nabla\big(\partial_u E(x,u)\big)\Big) + \mathbb{K}(x,u)\partial_u E(x,u).$$

We provide more explicit examples for $\mathcal{E}$ and $\mathcal{K}$ in Section 2.1.

In Chapter 3 we provide conditions on the system $(X, \mathcal{E}, \mathcal{G})$ such that the driving functional $\mathcal{E} : X \to \mathbb{R}_\infty$ is *geodesically λ-convex* with respect to the metric tensor $\mathcal{G} = \mathcal{K}^{-1}$. In particular, given a metric tensor $\mathcal{G}$ we can define a distance $\mathsf{d}_\mathcal{G} : X \times X \to [0, \infty]$ as the infimum of the *action functional* over all connecting curves, i.e.,

$$\mathsf{d}_\mathcal{G}(u_0, u_1)^2 = \inf\left\{ \textstyle\int_0^1 \langle \mathcal{G}(\gamma)\gamma', \gamma'\rangle\,\mathrm{d}s \;:\; \gamma \in \boldsymbol{C}(u_0, u_1) \right\},$$

where $\boldsymbol{C}(u_0, u_1)$ denotes the set of (sufficiently smooth) curves $\gamma : [0,1] \to X$ connecting u_0 and u_1, i.e., $\gamma(0) = u_0$ and $\gamma(1) = u_1$. Now, *(constant speed) geodesic curves* $\gamma : [0,1] \to X$ can be characterized as length minimizing curves, i.e.,

$$0 \leq s < t \leq 1 : \qquad \mathsf{d}_\mathcal{G}\big(\gamma(s), \gamma(t)\big) = |t{-}s|\, \mathsf{d}_\mathcal{G}\big(\gamma(0), \gamma(1)\big).$$

A functional $\mathcal{E} : X \to \mathbb{R}_\infty$ is called geodesic λ-convex, with respect to the distance $\mathsf{d}_\mathcal{G}$, if for all geodesic curves $\gamma : [0,1] \to X$ it satisfies the estimate

$$\mathcal{E}\big(\gamma(s)\big) \leq (1{-}s)\mathcal{E}\big(\gamma(0)\big) + s\mathcal{E}\big(\gamma(1)\big) - \lambda\frac{s(1{-}s)}{2}\mathsf{d}_\mathcal{G}\big(\gamma(0), \gamma(1)\big)^2.$$

The study of geodesic convex functionals in the context of optimal transport and the Wasserstein distance was started by McCann in [McC97] (there called displacement convex) and is studied extensively since then, see e.g. [OtW05, AGS05, DaS10, CL*10]. Geodesic λ-convex gradient structures $(X, \mathcal{E}, \mathcal{G})$ enjoy a number of nice properties (see Section 3.2.3 and [AGS05, Chapter 4]). The most important and most beneficial of them is the equivalent and purely metric description of the evolution of the gradient flow in terms of an *evolutionary variational inequality* (EVI_λ)

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\,\mathsf{d}_\mathcal{G}\big(u(t), w\big)^2 + \frac{\lambda}{2}\mathsf{d}_\mathcal{G}\big(u(t), w\big)^2 + \mathcal{E}\big(u(t)\big) \leq \mathcal{E}(w), \quad \forall\, w \in X,\ t > 0.$$

In order to establish the geodesic λ-convexity of $\mathcal{E}$ many methods rely on the characterization of geodesic curves and use tools from the theory of optimal transport and the Monge-Kantorovich formulation. However, the application of these tools is not possible for reaction-diffusion systems since in general the total mass is *not* conserved. Instead, our result, stated in Section 3.2, relies on a differential characterization of geodesic λ-convexity, which generalizes the approach of Daneri & Savaré [DaS08] (see also [OtW05]). These characterization involves the quadratic form

$$\xi \mapsto \langle \xi, \mathcal{M}(u)\xi \rangle \overset{\text{def}}{=} \langle \xi, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\xi \rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi \rangle,$$

which is in some sense the form induced by the metric Hessian of $\mathcal{E}$. The main result is that $\mathcal{E}$ is geodesically λ-convex if the crucial estimate

$$\langle \xi, \mathcal{M}(u)\xi \rangle \geq \lambda \langle \xi, \mathcal{K}(u)\xi \rangle \tag{2}$$

holds for all suitable u and ξ, see Proposition 3.2.7. Here, suitable means that u and ξ are smooth enough such that the quadratic form $\xi \mapsto \langle \xi, \mathcal{M}(u)\xi \rangle$ is well-defined. In particular, the main point of [DaS08] is that it suffices to establish the geodesic λ-convexity of $\mathcal{E}$ on a dense set, where all the calculations on functions can be done rigorously. Then the abstract theory allows us to extend the geodesic λ-convexity of the functional $\mathcal{E}$ to the closure of the domain of $\mathcal{E}$. The crucial point of the estimate (2) is that it permits for an arbitrary curve γ to control the change of its action while the whole curve γ evolves according to the semiflow induced by (1). In particular, for λ positive, the action is decaying. This allows us to probe the convexity of $\mathcal{E}$ (see [OtW05, DaS08]).

In Section 3.3 we collect possible applications of this abstract theory. First, we discuss simple reaction kinetics satisfying the detailed balance condition (see e.g. [GlG09]). This includes the case of general reversible Markov chains, see also [Maa11, Mie11a, ErM12]. Moreover, we treat partial differential equations or systems where the crucial estimate heavily relies on a well-chosen sequence of integrations by parts, where the occurring boundary integrals needs to be taken care of. We are able to generalize and complement the results in [Lis09] and [CL*10] for scalar diffusion equations (see Sections 3.3.2–3.3.3). Moreover, we present results for a scalar reaction-diffusion equation (Section 3.3.4), a linear reaction-diffusion system (Section 3.3.5), a one-dimensional drift-diffusion system (Section 3.3.6) and a multi-particle system with cross-diffusion of Stefan-Maxwell type (Section 3.3.7).

We emphasize that geodesic convexity is a strong structural property of a gradient system that is rather difficult to achieve. Our examples provide a first list of some nontrivial reaction-diffusion equations or systems that are geodesic λ-convex.

Finally, Chapter 4 deals with the application of variational methods to derive effective limit models for multiscale problems that exhibit a gradient structure $(X, \mathcal{E}, \mathcal{G})$. In particular, we contribute to the general endeavor of passing to the limit in evolution systems driven by functionals using variational methods such as *Γ-convergence*, see [SaS04, AM*12] for related applications to gradient structures and [MRS08] and [Mie08] for rate-independent and Hamiltonian systems, respectively. The main point in this investigation is that we only rely on the gradient structure of the systems.

We consider two model problems: the derivation of bulk/surface coupling for the Allen-Cahn equation in Section 4.1 and of an effective interface condition for a one-dimensional diffusion equation in Section 4.2.

In the first example we consider the following system of Allen-Cahn equations

$$\tau_{\mathrm{b}}\dot{u}_\varepsilon = A_{\mathrm{b}}\Delta u_\varepsilon - W_{\mathrm{b}}'(u_\varepsilon) \quad \text{in } \Omega, \quad \text{and} \quad \tau_\varepsilon \dot{u}_\varepsilon = A_\varepsilon \Delta u_\varepsilon - \tfrac{1}{\varepsilon} W_{\mathrm{s}}'(u_\varepsilon) \quad \text{in } \Sigma_\varepsilon,$$

posed in a fixed (bulk) domain Ω and boundary layer Σ_ε that surrounds Ω and shrinks to $\partial\Omega$ as the parameter ε tends to zero. We show, after a rescaling of the problem, that solutions of the above system converge to solutions of a limit system that describes the evolution of the bulk system coupled to an evolutionary system on the boundary $\partial\Omega$. Notably, we obtain as a limit the system which was recently studied by Sprekels and Wu in [SpW10].

For the limit passage we exploit De Giorgi's (Ψ, Ψ^*)-formulation which can be written in the integral form

$$\forall\, t \in [0,T]: \quad \mathcal{E}_\varepsilon\big(u_\varepsilon(t)\big) + \textstyle\int_0^t \Psi_\varepsilon(u_\varepsilon;\dot{u}_\varepsilon) + \Psi_\varepsilon^*\big(u_\varepsilon; -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon)\big)\,\mathrm{d}s = \mathcal{E}_\varepsilon\big(u_\varepsilon(0)\big),$$

with $\Psi_\varepsilon(u;\dot{u}) = \frac{1}{2}\langle \mathcal{G}_\varepsilon(u)\dot{u},\dot{u}\rangle$ and $\Psi_\varepsilon^*(u;\xi) = \frac{1}{2}\langle \xi, \mathcal{K}_\varepsilon(u)\xi\rangle$ denoting the dissipation and dual dissipation functionals, respectively. In particular, we adapt the ideas of Sandier & Serfaty in [SaS04] where an approach to prove the convergence of gradient flows for Γ-converging energy functionals was presented. This approach is similar to the concept of mutual recovery sequences introduced by Mielke, Roubíček and Stefanelli in [MRS08] and connects the convergence of the dissipation and energy functionals.

In the second example we consider the one-dimensional diffusion equation, i.e.,

$$\dot{u}_\varepsilon = (a_\varepsilon(x)u_\varepsilon')' \quad \text{in } \Omega \overset{\mathrm{def}}{=} \,]-\tfrac{1}{2},\tfrac{1}{2}[,$$

where the diffusion coefficient a_ε is of order ε in the small interval $]-\frac{\varepsilon}{2},\frac{\varepsilon}{2}[$. This equation is the gradient flow with respect to the logarithmic free energy and Wasserstein-type Onsager operator with spatial dependent diffusion coefficient a_ε.

We show that the solutions u_ε converge to a solution of the following limit diffusion equation with interface condition for $x = 0$

$$\dot{u} = \delta u'' \quad \text{in }]-\tfrac{1}{2},0[\, \cup\,]0,\tfrac{1}{2}[, \quad \text{coupled to} \quad \delta u_-' = k(u_+ - u_-) = \delta u_+' \quad \text{in } \{0\},$$

which is a simple bulk/interface system covered by Glitzky and Mielke in [GlM13] (see also [Mie13]). There, it was shown that the limit equation can be formulated in terms of the gradient system

$$\mathcal{E}(u) = \textstyle\int_\Omega u\log u\,\mathrm{d}x, \quad \text{and} \quad \Psi^*(u,\xi) = \frac{1}{2}\int_{\Omega\setminus\{0\}} \delta u|\xi'|^2\,\mathrm{d}x + \frac{k}{2}\Lambda(u_+,u_-)(\xi_+ - \xi_-)^2,$$

where $\Lambda(a,b) = (a-b)/(\log a - \log b)$ denotes the logarithmic mean of a and b. In particular, the proof of the convergence also uses a rescaling of the interface layer and follows the ideas in [AM*12]. There, a similar limit was discussed, namely the passage from diffusion in a one-dimensional Fokker-Planck equation to (linear) reaction.

2 Onsager operators and reaction-diffusion systems

In several papers by Otto (see [JKO98, Ott98, Ott01]) it was shown that certain diffusion problems can be interpreted as gradient flows with respect to the free energy or relative entropy and the Wasserstein distance. In [Mie11b] it was shown that general reaction-diffusion systems, with reactions satisfying the detailed balance condition, can be written as a gradient system with respect to the relative entropy.

In an abstract context we understand a gradient system as a triple $(X, \mathcal{E}, \mathcal{G})$ where X is the state space containing the states $u \in X$. For simplicity we assume that X is a reflexive Banach space with dual X^*. The driving functional $\mathcal{E} : X \to \mathbb{R}_\infty \stackrel{\text{def}}{=} \mathbb{R} \cup \{\infty\}$ is assumed to be differentiable (in a suitable way) such that the *potential restoring force* is given by $-\mathrm{D}\mathcal{E}(u) \in X^*$. The third ingredient is a metric tensor $\mathcal{G}$, i.e., $\mathcal{G}(u) : X \to X^*$ is linear, symmetric and positive (semi-)definite.

The gradient flow associated with $(X, \mathcal{E}, \mathcal{G})$ is the (abstract) *force balance*

$$\mathcal{G}(u)\dot{u} = -\mathrm{D}\mathcal{E}(u) \qquad \Longleftrightarrow \qquad \dot{u} = -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u) \stackrel{\text{def}}{=} -\nabla_{\mathcal{G}}\mathcal{E}(u), \tag{2.1}$$

where we recall that the "gradient" $\nabla_{\mathcal{G}}\mathcal{E}$ of the functional $\mathcal{E}$ is an element of X (in contrast to the differential $\mathrm{D}\mathcal{E}(u) \in X^*$) and is calculated in terms of $\mathcal{K}(u) = \mathcal{G}(u)^{-1}$. We call this equation an abstract force balance, since $\mathcal{G}(u)\dot{u}$ can be seen as a viscous force arising from the motion of u. We call the linear, symmetric and positive semidefinite operator $\mathcal{K}(u) : X^* \to X$ the *Onsager operator* and the corresponding triple $(X, \mathcal{E}, \mathcal{K})$ *Onsager system.*

Since we are mainly interested in reaction-diffusion systems we consider (vectors of) densities $\boldsymbol{u} : \Omega \to \,]0, \infty[^I$ of diffusive species $A_1, \ldots, A_I$. Moreover, the driving functional of the evolution $\mathcal{E}$ shall be of the form

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(x, \boldsymbol{u}(x))\,\mathrm{d}x,$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain and $E : \overline{\Omega} \times \mathbb{R}^I \to \mathbb{R}$ is a sufficiently smooth energy density. It was shown in [Mie11b] that for a wide class of reaction-diffusion systems gradient, or equivalently, Onsager structures can be specified.

A major advantage of the Onsager form is its flexibility in modeling: Quite often differential equations are written in rate form where the vector field is additively decomposed into different physical phenomena. This additive split can be also used for the Onsager operator, as long as all the different effects are driven by the same functional. In particular, since we are interested in reaction-diffusion systems we shall consider the decomposition of $\mathcal{K}$ into a *diffusive* and a *reaction* part, namely $\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = \mathcal{K}_{\text{diff}}(\boldsymbol{u})\boldsymbol{\xi} + \mathcal{K}_{\text{react}}(\boldsymbol{u})\boldsymbol{\xi}$. Here, $\boldsymbol{\xi}$ is the *thermodynamically conjugated force* being dual to the rate $\dot{\boldsymbol{u}}$.

Following the Wasserstein approach to diffusion introduced by OTTO in [JKO98, Ott01] – also called OTTO's formalism – we define the diffusion part as

$$\mathcal{K}_{\text{diff}}(\boldsymbol{u})\boldsymbol{\xi} = -\text{div}\big(\mathbb{M}(x, \boldsymbol{u})\nabla\boldsymbol{\xi}\big) \tag{2.2a}$$

with $\mathbb{M}(x, \boldsymbol{u}) \in \text{Lin}(\mathbb{R}^{I\times d}, \mathbb{R}^{I\times d})$ being a symmetric and positive semidefinite fourth order tensor. In turn, the reaction part $\mathcal{K}_{\text{react}}$ is given by a symmetric and positive semidefinite matrix $\mathbb{K}(x, \boldsymbol{u}) \in \mathbb{R}^{I\times I}$, i.e.,

$$\mathcal{K}_{\text{react}}(\boldsymbol{u})\boldsymbol{\xi} = \mathbb{K}(x, \boldsymbol{u})\boldsymbol{\xi}. \tag{2.2b}$$

Using these definitions the abstract force balance in (2.1), which describes the evolution of the densities $\boldsymbol{u}$, takes the form

$$\dot{\boldsymbol{u}} = -\text{div}\Big(\mathbb{M}(x, \boldsymbol{u})\nabla\big(\mathrm{D}_{\boldsymbol{u}}E(x, \boldsymbol{u})\big)\Big) + \mathbb{K}(x, \boldsymbol{u})\mathrm{D}_{\boldsymbol{u}}E(x, \boldsymbol{u}), \tag{2.3}$$

subjected to the no-flux boundary condition $\mathbb{M}(x, \boldsymbol{u})\nabla\big(\mathrm{D}_{\boldsymbol{u}}E(x, \boldsymbol{u})\big) \cdot \nu(x) = 0$ for $x \in \partial\Omega$.

The symmetry of the tensor $\mathcal{K}(\boldsymbol{u})$ allows us to define the *dual dissipation potential*

$$\Psi^*(\boldsymbol{u}; \boldsymbol{\xi}) = \frac{1}{2}\langle\boldsymbol{\xi}, \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \frac{1}{2}\int_\Omega \nabla\boldsymbol{\xi}\cdot\mathbb{M}(x, \boldsymbol{u})\nabla\boldsymbol{\xi} + \boldsymbol{\xi}\cdot\mathbb{K}(x, \boldsymbol{u})\boldsymbol{\xi}\,\mathrm{d}x.$$

We call Ψ^* the dual dissipation potential since it is the Legendre transform of the *dissipation potential* $\Psi : (\boldsymbol{u}, \dot{\boldsymbol{u}}) \mapsto \frac{1}{2}\langle\mathcal{G}(\boldsymbol{u})\dot{\boldsymbol{u}}, \dot{\boldsymbol{u}}\rangle$, i.e., we have

$$\begin{aligned}\Psi(\boldsymbol{u}; \boldsymbol{v}) &= \sup\left\{\langle\boldsymbol{\xi}, \boldsymbol{v}\rangle - \Psi^*(\boldsymbol{u}, \boldsymbol{\xi}) \;:\; \boldsymbol{\xi} \in X^*\right\},\\ \Psi^*(\boldsymbol{u}; \boldsymbol{\xi}) &= \sup\left\{\langle\boldsymbol{\xi}, \boldsymbol{v}\rangle - \Psi(\boldsymbol{u}, \boldsymbol{v}) \;:\; \boldsymbol{v} \in X\right\}.\end{aligned}$$

Using the classical Legendre equivalences for convex functionals $\mathcal{J} : X \to [0, \infty]$, namely

$$\boldsymbol{\xi} \in \partial\mathcal{J}(\boldsymbol{v}) \quad\iff\quad \boldsymbol{v} \in \partial\mathcal{J}^*(\boldsymbol{\xi}) \quad\iff\quad \mathcal{J}(\boldsymbol{v}) + \mathcal{J}^*(\boldsymbol{\xi}) = \langle\boldsymbol{\xi}, \boldsymbol{v}\rangle$$

and the chain rule for $t \mapsto \mathcal{E}(\boldsymbol{u}(t))$ we find the formulation equivalent to (2.1)

$$\mathcal{E}\big(\boldsymbol{u}(0)\big) - \mathcal{E}\big(\boldsymbol{u}(t)\big) = \int_0^t \Psi(\boldsymbol{u}; \dot{\boldsymbol{u}}) + \Psi^*\big(\boldsymbol{u}; -\mathrm{D}\mathcal{E}(\boldsymbol{u})\big)\,\mathrm{d}s. \tag{2.4}$$

The crucial point is that although this so-called (Ψ, Ψ^*)-*formulation* is a scalar equation it already describes the dynamics of the system completely. In the theory of gradient flows in metric spaces this formulation is better known as the *De Giorgi principle*, and solutions are called *curves of maximal slope* (see [DMT80, AGS05, DaS10]). In the following, we will also call $(X, \mathcal{E}, \mathcal{K})$ and $(X, \mathcal{E}, \Psi)$ "gradient system".

In particular, the formulation in (2.4) allows us to apply tools from the calculus of variations such as Γ- and Mosco convergence (see Chapter 4).

2.1 Examples

We conclude this chapter by providing some illustrative examples of gradient systems, which we will revisit in the subsequent chapters, e.g., when discussing geodesic λ-convexity in Chapter 3. We refer to [Mie11b], [GlM13] and [Mie13] for a more detailed discussion of these systems.

2.1.1 Allen-Cahn equation and Cahn-Hilliard equation

Probably the most well-known examples of gradient flows are the Allen-Cahn and Cahn-Hilliard equations which are given in terms of the free energy $\mathcal{E}(u) = \int_\Omega \frac{\alpha}{2}|\nabla u|^2 + W(u)\,\mathrm{d}x$. The Allen-Cahn equation takes the form

$$\dot{u} = -k_{\mathrm{AC}}\mathrm{D}\mathcal{E}(u) = -k_{\mathrm{AC}}\big(-\mathrm{div}(\alpha\nabla u) + W'(u)\big).$$

In particular, the Onsager operator is the multiplication operator $\mathcal{K}_{\mathrm{AC}}(u)\xi = k_{\mathrm{AC}}\xi$, and the dual dissipation potential is given via $\Psi^*_{\mathrm{AC}}(\xi) = \int_\Omega \frac{k_{\mathrm{AC}}}{2}|\xi|^2\,\mathrm{d}x$. We will return to this particular system in Chapter 4 when we discuss the application of variational methods such as Γ-convergence to derive asymptotic models for bulk/surface coupling.

In contrast, the Cahn-Hilliard equation for the (conserved) phase parameter φ is a parabolic equation of fourth order and reads

$$\dot{\varphi} = -\mathcal{K}_{\mathrm{CH}}(\varphi)\mathrm{D}\mathcal{E}(\varphi) = -\mathrm{div}\Big(\mathbb{M}(\varphi)\nabla\big(-\mathrm{div}(\alpha\nabla\varphi) + W'(\varphi)\big)\Big).$$

Hence, the Onsager operator is the differential operator $\mathcal{K}_{\mathrm{CH}}(\varphi)\xi = -\mathrm{div}(\mathbb{M}(\varphi)\nabla\xi)$, see also [LMS12]. Note that the evolution leaves the averages $\int_\Omega \varphi(t,x)\,\mathrm{d}x$ constant in time. This follows from the general property of $\mathcal{K}_{\mathrm{CH}}$ that for $\xi = c = \mathrm{const}$ we have $\mathcal{K}_{\mathrm{CH}}(\varphi)c = 0$.

2.1.2 Chemical reaction kinetics of mass-action type

Pure chemical reaction systems are ODE systems $\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u})$, where often the right-hand side is written in terms of polynomials associated to the reaction kinetics. It was observed in [Mie11b] that under the assumption of detailed balance (also called reversibility) such systems have a gradient structure with the relative entropy

$$E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i)$$

as the driving functional, where the $w_i > 0$ denote fixed reference densities. We assume that there are R reactions of mass-action type (cf. e.g. [DeM84, GiM04, KjB08]) between the species $A_1, \dots, A_I$ written as

$$\alpha_1^r A_1 + \cdots + \alpha_I^r A_I \underset{k_r^{\mathrm{bw}}}{\overset{k_r^{\mathrm{fw}}}{\rightleftharpoons}} \beta_1^r A_1 + \cdots + \beta_I^r A_I \qquad r = 1, \dots, R,$$

where $k_r^{\mathrm{bw}} > 0$ and $k_r^{\mathrm{fw}} > 0$ are the backward and forward reaction rates that may also depend on the densities of the species. The vectors $\boldsymbol{\alpha}^r,\ \boldsymbol{\beta}^r \in \mathbb{N}_0^I$ contain the stoichiometric coefficients of the rth reaction.

The associated reaction system for the densities (in a spatially homogeneous system, where diffusion can be neglected) reads

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) \overset{\mathrm{def}}{=} -\sum_{r=1}^{R} \Big(k_r^{\mathrm{fw}}(\boldsymbol{u})\, \boldsymbol{u}^{\boldsymbol{\alpha}^r} - k_r^{\mathrm{bw}}(\boldsymbol{u})\, \boldsymbol{u}^{\boldsymbol{\beta}^r}\Big)\Big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\Big), \tag{2.5}$$

where we use the monomial notation $\boldsymbol{u}^{\boldsymbol{\alpha}} = u_1^{\alpha_1} \cdots u_I^{\alpha_I} \in \mathbb{R}$.

The main assumption to obtain a gradient structure is that of *detailed balance*, which means that there exists a reference density vector $\boldsymbol{w}$ such that all R reactions are balanced individually, namely for all $r = 1, \ldots, R$ and all $\boldsymbol{u} \in \,]0,\infty[^I$

$$\text{there exists } \boldsymbol{w} \in \,]0,\infty[^I \text{ such that } k_r^*(\boldsymbol{u}) \overset{\mathrm{def}}{=} k_r^{\mathrm{fw}}(\boldsymbol{u})\boldsymbol{w}^{\boldsymbol{\alpha}^r} = k_r^{\mathrm{bw}}(\boldsymbol{u})\boldsymbol{w}^{\boldsymbol{\beta}^r}.$$

As in [Mie11b] we now define the Onsager matrix

$$\mathbb{K}(\boldsymbol{u}) = \sum_{r=1}^{R} k_r^*(\boldsymbol{u})\Lambda\big(\tfrac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}}, \tfrac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}}\big)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big)\otimes\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \text{ with } \Lambda(a,b) = \frac{a-b}{\log a - \log b}$$

and find that the reaction system (2.5) takes the form

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) = -\mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u}).$$

This follows easily by using the definition of the logarithmic mean Λ and the calculation rules for logarithms, namely, for $\mathbf{log}\,\boldsymbol{u} = (\log u_i)_{i=1,\ldots,I}$ we compute

$$\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big)\cdot\big(\mathbf{log}\,\boldsymbol{u} - \mathbf{log}\,\boldsymbol{w}\big) = \log\big(\boldsymbol{u}^{\boldsymbol{\alpha}^r}/\boldsymbol{w}^{\boldsymbol{\alpha}^r}\big) - \log\big(\boldsymbol{u}^{\boldsymbol{\beta}^r}/\boldsymbol{w}^{\boldsymbol{\beta}^r}\big).$$

2.1.3 Diffusion equations

For the gradient structure of diffusion systems $\dot{\boldsymbol{u}} = \mathrm{div}\big(\mathbb{M}(\boldsymbol{u})\nabla\boldsymbol{u}\big)$ one might be tempted to use a functional involving the gradient $\nabla\boldsymbol{u}$. However, we have to use the relative entropy as a driving functional, because we must use the same functional for modeling the reactions. Hence, we adopt the Wasserstein approach to diffusion introduced by Otto and coauthors: For $\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(u)\,\mathrm{d}x$ the diffusion system will take the form $\dot{\boldsymbol{u}} = -\mathcal{K}_{\mathrm{diff}}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u})$ with an Onsager operator $\mathcal{K}_{\mathrm{diff}}$ given via

$$\mathcal{K}_{\mathrm{diff}}(\boldsymbol{u})\boldsymbol{\xi} = -\mathrm{div}\big(\widetilde{\mathbb{M}}(\boldsymbol{u})\nabla\boldsymbol{\xi}\big),$$

where $\widetilde{\mathbb{M}}(\boldsymbol{u}) : \mathbb{R}^{I\times d} \to \mathbb{R}^{I\times d}$ is a symmetric and positive semi-definite tensor of order four such that $\mathbb{M}(\boldsymbol{u}) = \widetilde{\mathbb{M}}(u)\mathrm{D}^2E(\boldsymbol{u})$. Hence, the Onsager system leads to the diffusion system

$$\dot{\boldsymbol{u}} = \mathrm{div}\big(\widetilde{\mathbb{M}}(\boldsymbol{u})\nabla\mathrm{D}E(\boldsymbol{u})\big) = \mathrm{div}\big(\mathbb{M}(\boldsymbol{u})\nabla\boldsymbol{u}\big).$$

We emphasize that $\widetilde{\mathbb{M}}$ has to be symmetric by ONSAGER's symmetry relations, which leads to nonsymmetric operators $\mathbb{M}$, if there is cross-diffusion (see e.g. [Mie13, Sect. 3.2.1] for a simple example).

2.1.4 Coupling diffusion and reaction

Now, we consider coupled reaction-diffusion systems. The driving functional for the evolution is the total relative entropy $\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(\boldsymbol{u})\,\mathrm{d}x$. The Onsager operator is given by the sum $\mathcal{K}(\boldsymbol{u}) = \mathcal{K}_{\mathrm{diff}}(\boldsymbol{u}) + \mathcal{K}_{\mathrm{react}}(\boldsymbol{u})$ with $\mathcal{K}_{\mathrm{diff}}$ and $\mathcal{K}_{\mathrm{react}}$ as in (2.2). Hence, the coupled system reads

$$\dot{\boldsymbol{u}} = \mathrm{div}\big(\widetilde{\mathbb{M}}(\boldsymbol{u})\nabla \mathrm{D}E(\boldsymbol{u})\big) + \mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u}) = \mathrm{div}\big(\mathbb{M}(\boldsymbol{u})\nabla \boldsymbol{u}\big) + \boldsymbol{R}(\boldsymbol{u}),$$

where $\mathbb{M}(\boldsymbol{u}) = \widetilde{\mathbb{M}}(\boldsymbol{u})\mathrm{D}^2E(\boldsymbol{u})$ and $\boldsymbol{R}(\boldsymbol{u}) = \mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u})$.

As an example for a reaction-diffusion system we consider the quaternary system studied in [DF*07, DeF08], namely the evolution of a mixture of diffusive species A_1, A_2, A_3 and A_4 in a bounded domain Ω undergoing a reversible reaction of the type

$$A_1 + A_2 \underset{k^{\mathrm{bw}}}{\overset{k^{\mathrm{fw}}}{\rightleftharpoons}} A_3 + A_4. \tag{2.6}$$

For the density vector $\boldsymbol{u} = (u_1, u_2, u_3, u_4)$ we introduce the free energy functional

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega \sum_{i=1}^{4} u_i \log(u_i/w_i)\,\mathrm{d}x.$$

For simplicity we assume that $k^{\mathrm{fw}} = k^{\mathrm{bw}} = 1$ and can take $w_i = 1$. We have the stoichiometric vectors $\boldsymbol{\alpha} = (1,1,0,0)$, $\boldsymbol{\beta} = (0,0,1,1)$ and thus

$$\mathbb{K}(u_1, u_2, u_3, u_4) = \Lambda\big(u_1u_2\,,\, u_3u_4\big)\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix}.$$

With the tensor $\mathbb{M}(\boldsymbol{u}) = \mathrm{diag}(\delta_1u_1, \ldots, \delta_4u_4)$ we define the corresponding Onsager operator $\mathcal{K}_{\mathrm{diff}}$ which leads to the reaction-diffusion system

$$\dot{\boldsymbol{u}} = \mathrm{div}\big(\mathbb{D}\nabla\boldsymbol{u}\big) - (u_1u_2 - u_3u_4)(\boldsymbol{\alpha}-\boldsymbol{\beta}), \quad \text{where } \mathbb{D} = \mathrm{diag}(\delta_1, \ldots, \delta_4).$$

In fact, many reaction-diffusion systems studied in the literature (including semiconductor models involving an elliptic equation for the electrostatic potential), see e.g. [GlH05, DeF06, DeF07, Gli09, BoP11], have the structure developed above. But except for the recent work [Mie11b, GlM13, Mie13], the gradient structure was not displayed and used explicitly, only the *Liapunov property* of the free energy $\mathcal{E}$ was exploited for deriving a priori estimates.

2.1.5 Drift-reaction-diffusion equations

We close this section by considering a drift-diffusion system coming from the theory of semiconductor devices. More precisely, we treat a simple semiconductor model related to the *van Roosbroeck system* (see [GaG86]). Here, we additionally need to take into account that the electric charge of the species generates an *electric potential*, whose electric field creates drift forces proportional to the charges of the species. We recite here briefly the results of [Mie11b, Sect. 4] and refer to latter for the full discussion. Moreover, we refer to [GlM13] for drift-diffusion systems exhibiting bulk-interface interaction.

The system's state is described by the electron and hole densities $n : \Omega \to]0, \infty[$ and $p : \Omega \to]0, \infty[$, respectively. The charged species generate an electrostatic potential $\phi_{n,p}$ being the unique solution of the linear potential equation

$$-\mathrm{div}(\varepsilon \nabla \phi) = \delta + q_n n + q_p p \quad \text{in } \Omega, \qquad \phi = \phi_{\mathrm{Dir}} \quad \text{on } \Gamma_{\mathrm{Dir}} \subset \partial\Omega, \tag{2.7a}$$

where $\delta : \Omega \to \mathbb{R}$ is a given doping profile and $q_n = -1$ and $q_p = 1$ are the charge numbers with opposite sign. The evolution of the densities n, p is governed by diffusion, drift with respect to the electric field $\nabla\phi_{n,p}$, and recombination according to simple creation-annihilation reactions for electron-hole pairs (radiative recombination), namely

$$A_n + A_p \rightleftharpoons \emptyset, \quad \text{i.e., } \boldsymbol{\alpha} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \boldsymbol{\beta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

With mobilities $\mu_n(n, p), \mu_p(n, p) > 0$ and reaction rate $\kappa(n, p) > 0$ the drift-diffusion system reads

$$\begin{aligned} \dot{n} &= \mathrm{div}\big(\mu_n(n, p)(\nabla n + q_n n \nabla \phi_{n,p})\big) - \kappa(n, p)(np-1), \\ \dot{p} &= \mathrm{div}\big(\mu_p\,(n, p)(\nabla p + q_p\, p \nabla \phi_{n,p})\big) - \kappa(n, p)(np-1). \end{aligned} \tag{2.7b}$$

For establishing a gradient structure we define the functional $\mathcal{E}$ as the sum of electrostatic and free energy:

$$\mathcal{E}(n, p) = \int_\Omega \frac{1}{2}|\nabla\phi_{n,p}|^2 + n(\log n - 1) + p(\log p - 1)\,\mathrm{d}x.$$

The thermodynamic conjugated forces, also called quasi-Fermi potentials or electrochemical potentials, read

$$\mathrm{D}_n\mathcal{E}(n, p) = \log n + q_n\phi_{n,p} \quad \text{and} \quad \mathrm{D}_p\mathcal{E}(n, p) = \log p + q_p\phi_{n,p}.$$

Here we used that $\phi_{n,p}$ solves (2.7a) and depends affinely on n and p. The Onsager operator $\mathcal{K}(n, p)$ takes the form

$$\mathcal{K}(n, p)\begin{pmatrix} \xi_n \\ \xi_p \end{pmatrix} = \begin{pmatrix} -\mathrm{div}(\mu_n n \nabla \xi_n) \\ -\mathrm{div}\mu_p(p\nabla\xi_p) \end{pmatrix} + \kappa(n, p)\Lambda(np, 1)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \xi_n \\ \xi_p \end{pmatrix}.$$

Thus, again we have two Wasserstein terms for the electrochemical potentials coupled with a reaction term. We immediately find that for $q_n = -q_p$ (opposite charges of electron and holes) it holds that $\binom{q_n}{q_p} \in \operatorname{Ker} \mathcal{K}(n,p)$. This means, that the total charge $\mathcal{Q}(n,p) = \int_\Omega \delta + q_n n + q_p p \, \mathrm{d}x$ is a conserved quantity, i.e., $\frac{\mathrm{d}\mathcal{Q}(n,p)}{\mathrm{d}t} = 0$. Moreover, using that

$$-\mathcal{K}(n,p)\mathrm{D}\mathcal{E}(n,p) = \begin{pmatrix} \operatorname{div}\big(\mu_n n \nabla(\log n + q_n \phi_{n,p})\big) - \kappa\Lambda(np,1)\log(np) \\ \operatorname{div}\big(\mu_p p \nabla(\log p + q_p \phi_{n,p})\big) - \kappa\Lambda(np,1)\log(np) \end{pmatrix}$$

we see that $\binom{\dot n}{\dot p} = -\mathcal{K}(n,p)\mathrm{D}\mathcal{E}(n,p)$ is the desired Onsager structure of the van Roosbroeck system (2.7).

A similar gradient system with only one species was considered in [AmS08], namely

$$\dot u = \operatorname{div}\big(u \nabla \Phi_u\big), \; -\Delta\Phi_u + \Phi_u = u \text{ in } \Omega, \quad \nabla u \cdot \nu = 0, \; \Phi = 1 \text{ on } \partial\Omega.$$

It is a gradient system for the energy $\mathcal{E}(u) = \int_\Omega u + \frac{1}{2}|\nabla \Phi_u|^2 + \frac{1}{2}|\Phi_u - 1|^2 \, \mathrm{d}x$ and the Wasserstein operator $\mathcal{K}(u)\xi = -\operatorname{div}(u\nabla\xi)$.

3 Geodesic convexity for gradient systems

The aim of this chapter is to provide conditions on gradient systems $(X, \mathcal{E}, \mathcal{G})$ introduced in the previous chapter such that the driving functional $\mathcal{E}$ is *geodesically λ-convex* with respect to the metric $\mathcal{G} = \mathcal{K}^{-1}$. Geodesic λ-convexity of $\mathcal{E}$ with respect to $\mathcal{G}$ means that there exists a $\lambda \in \mathbb{R}$ such that for each (constant speed) geodesic curves $\gamma : [0,1] \to X$ (see (3.12) for the definition) and for each $s \in [0,1]$

$$\mathcal{E}(\gamma(s)) \le (1-s)\mathcal{E}(\gamma(0)) + s\mathcal{E}(\gamma(1)) - \lambda\frac{s(1-s)}{2}\mathsf{d}_{\mathcal{K}}(\gamma(0),\gamma(1))^2. \tag{3.1}$$

Here, $\mathsf{d}_{\mathcal{K}} : X \times X \to [0,\infty]$ denotes the distance induced by the metric tensor $\mathcal{G}$ and is defined as the infimum of the *action functional* $\mathcal{A}(\gamma, \gamma') = \langle \mathcal{G}(\gamma)\gamma', \gamma' \rangle$ over all connecting curves $\gamma : [0,1] \to X$ (see (3.10)), where γ' denotes the derivative with respect to the arclength parameter s.

The study of geodesic λ-convexity for scalar drift-diffusion equations given by

$$\mathcal{E}(u) = \int_\Omega E(u) + uV(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{K}(u)\xi = -\mathrm{div}(\mu(u)\nabla\xi), \tag{3.2}$$

was initiated by McCann in [McC97] (there called *displacement convexity*) and is studied extensively since then, see e.g. [Stu05, OtW05, AGS05, DaS10, CL*10]. An essential tool in this theory is the characterization of the geodesic curves in terms of mass transportation and the optimal transport problem of Monge-Kantorovich type.

Presently, such a method is not available for systems of equations or for scalar equations with reaction terms, which destroy the conservation of mass. Instead, the results in [LiM12], which are presented in this chapter, rely on a differential characterization of geodesic λ-convexity developed by Daneri and Savaré in [DaS08].

In Section 3.2 we provide an abstract version of the theory developed by Daneri and Savaré in [DaS08]. We mainly address the abstract framework and present the estimates to obtain concrete convexity properties, while the functional analytic aspects as well as the full framework in terms of complete metric spaces are postponed to subsequent work. Moreover, we assume that our evolutionary system

$$\dot{u} = -\mathcal{F}(u) \stackrel{\text{def}}{=} -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u) \tag{3.3}$$

generates a suitable smooth local semiflow on a scale of Banach spaces $Z \subset Y \subset H$ with dense embeddings, see Section 3.2 for the details. The main characterization of geodesically convex gradient systems $(\mathcal{X}, \mathcal{E}, \mathcal{K})$ involves the quadratic form

$$\xi \mapsto \langle \xi, \mathcal{M}(u)\xi \rangle \stackrel{\text{def}}{=} \langle \xi, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\xi \rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi \rangle,$$

which can be seen as the form induced by the metric Hessian of $\mathcal{E}$. The main result is that $\mathcal{E}$ is geodesically λ-convex if the estimate

$$\langle \xi, \mathcal{M}(u)\xi \rangle \geq \lambda \langle \xi, \mathcal{K}(u)\xi \rangle \tag{3.4}$$

holds for all suitable u and ξ, see Proposition 3.2.7. Thus, the maximal λ satisfying this estimate is characterized by

$$\lambda^*_{\mathcal{E},\mathcal{K}} = \inf \left\{ \frac{\langle \xi, \mathcal{M}(u)\xi \rangle}{\langle \xi, \mathcal{K}(u)\xi \rangle} \;:\; u, \xi \text{ suitable} \right\}.$$

In particular, for flat geometries $\mathcal{K}(u) \equiv \mathcal{K}$ we recover the standard conditions

$$\langle \mathcal{K}\xi, \mathrm{D}^2\mathcal{E}(u)\mathcal{K}\xi \rangle \geq \lambda \langle \xi, \mathcal{K}\xi \rangle \qquad \text{or rather} \qquad \langle \mathrm{D}^2\mathcal{E}(u)v, v \rangle \geq \lambda \langle \mathcal{G}v, v \rangle.$$

Our proof is a straightforward generalization of the approach in [DaS08] which in turn is based on the evolutionary variational inequality (EVI_λ) given by

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathsf{d}_{\mathcal{K}}(u(t), w)^2 + \frac{\lambda}{2}\mathsf{d}_{\mathcal{K}}(u(t), w)^2 + \mathcal{E}(u(t)) \leq \mathcal{E}(w), \quad \forall\, w \in X,\ t > 0, \tag{3.5}$$

where $\frac{\mathrm{d}^+}{\mathrm{d}t} f(t) = \limsup_{\tau \downarrow 0} \frac{1}{\tau}(f(t+\tau) - f(t))$ is the right-upper Dini derivative. The idea is to use the semiflow induced by (3.3) – on a dense subset of X where all computations can be made rigorous – and the estimate in (3.4) to obtain (EVI_λ). Finally from (EVI_λ) we deduce (3.1) (see Theorem 3.2.2).

Let us emphasize from the very beginning that we assume throughout this chapter that the semiflow generated by (3.3) is given and has sufficient regularity properties (see Section 3.2.4).

In the main part of this chapter in Section 3.3 we collect possible applications of the abstract theory developed in Section 3.2. We stress that geodesic convexity is a strong structural property of a gradient system that is rather difficult to achieve, in particular with respect to distances $\mathsf{d}_{\mathcal{G}}$ that are associated with the Wasserstein metric. Our examples show that there are at least some nontrivial reaction-diffusion equations or systems that satisfy this beautiful property. First we discuss simple reaction kinetics satisfying the detailed balance conditions, i.e., ODE systems in the form

$$\dot{\boldsymbol{u}} = -F(\boldsymbol{u}) \stackrel{\text{def}}{=} -\mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u}), \quad \text{where } E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i).$$

This includes the case of general reversible Markov chains $\dot{\boldsymbol{u}} = Q\boldsymbol{u}$, where $Q \in \mathbb{R}^{I \times I}$ is a stochastic generator (intensity matrix), see also [Maa11, Mie11a, ErM12].

In the subsequent subsections we treat partial differential equations or systems where estimate (3.4) heavily relies on a well-chosen sequence of integrations by parts, where the occurring boundary integrals need to be taken care of. Here, we use the fact that for convex domains Ω and functions $\xi \in \mathrm{H}^3(\Omega)$ with $\nabla\xi \cdot \nu = 0$ on $\partial\Omega$, we have $\nabla\big(|\nabla\xi|^2\big) \cdot \nu \leq 0$ on $\partial\Omega$, Proposition 3.3.2. In Section 3.3.2 we give a lower bound for the geodesic convexity of $\mathcal{E}(u) = \int_\Omega u \log u \,\mathrm{d}x$ with respect to the inhomogeneous Wasserstein distance induced by $\mathcal{K}(u)\xi = -\mathrm{div}(\mu(x)u\nabla\xi)$, where $0 < \mu_0 \leq \mu \in \mathrm{W}^{2,\infty}(\Omega)$, thus generalizing results in [Lis09]. Theorem 3.3.3 provides a new result of geodesic convexity for $\mathcal{E}$ and $\mathcal{K}$ from (3.2), where the concave mobility $u \mapsto \mu(u)$ is allowed to be decreasing, i.e. $\mu'(u) < 0$, thus complementing results in [CL*10].

Sections 3.3.4 and 3.3.5 discuss problems with reactions, namely

$$\dot{u} = \Delta u - f(u) \quad \text{and} \quad \begin{pmatrix} \dot{u}_1 \\ \dot{u}_2 \end{pmatrix} = \begin{pmatrix} \delta\Delta u_1 \\ \delta\Delta u_2 \end{pmatrix} + k \begin{pmatrix} u_2 - u_1 \\ u_1 - u_2 \end{pmatrix}.$$

The first case with $f(u) = k(1-u)$ gives geodesic λ-convexity with $\lambda = \frac{1}{2}k$, while the second case gives geodesic 0-convexity. In Section 3.3.6 a one-dimensional drift-diffusion system with charged species is considered, where the nonlinear coupling occurs via the electrostatic potential. The final example discusses cross-diffusion of Stefan-Maxwell type for $\boldsymbol{u} = (u_1, ..., u_I)$ under the size-exclusion condition $u_1 + \cdots + u_I \equiv 1$ (see [Gri04]).

There are further interesting applications of gradient flows where methods based on geodesic convexity can be employed, even though the system under investigation may not be geodesically λ-convex, see e.g. the fourth order problems studied in [MMS09, GST09, CL*10]. Possible applications to viscoelasticity are discussed in [MOS12]. In [FiG10] a diffusion equation with Dirichlet boundary conditions, which leads to absorption, is investigated.

3.1 A formal derivation of the key estimate

Assuming that geodesic curves in the state space $\mathcal{X}$ are sufficiently smooth we can derive the crucial estimate in (3.4) using the following characterization of geodesic curves in terms of the Onsager operator $\mathcal{K}$ rather than of the Riemannian tensor $\mathcal{G}$. A geodesic curve $\gamma : [0,1] \to X$ satisfies the classical Lagrange equation

$$-\frac{\mathrm{d}}{\mathrm{d}s}\Big(\frac{\partial}{\partial\gamma'}\mathcal{L}(\gamma,\gamma')\Big) + \frac{\partial}{\partial\gamma}\mathcal{L}(\gamma,\gamma') = 0, \qquad \text{where } \mathcal{L}(\gamma,\gamma') = \frac{1}{2}\langle\mathcal{G}(\gamma)\gamma',\gamma'\rangle.$$

However, in the cases we are interested in $\mathcal{G}$ is only known implicitly, thus it is more convenient to use the Hamiltonian version of the Lagrange equation. Introducing the dual variable $\xi = \frac{\partial}{\partial\gamma'}\mathcal{L}(\gamma,\gamma') = \mathcal{G}(\gamma)\gamma'$ and the Hamiltonian $\mathcal{H}(\gamma,\xi) = \frac{1}{2}\langle\xi,\mathcal{K}(\gamma)\xi\rangle$ we obtain the equivalent system

$$\gamma' = \frac{\partial}{\partial\xi}\mathcal{H}(\gamma,\xi) = \mathcal{K}(\gamma)\xi, \qquad \xi' = -\frac{\partial}{\partial\gamma}\mathcal{H}(\gamma,\xi) = -\frac{1}{2}\langle\xi,\mathrm{D}\mathcal{K}(\gamma)[\Box]\xi\rangle, \tag{3.6}$$

where $b = \langle \xi, \mathrm{D}\mathcal{K}(\gamma)[\square]\xi\rangle$ denotes the vector defined via $\langle b, v\rangle = \langle \xi, \mathrm{D}\mathcal{K}(\gamma)[v]\xi\rangle$. Now, geodesic λ-convexity of a functional $\mathcal{E} : X \to \mathbb{R}$ can be easily characterized by asking that for $s \in [0,1]$ the composition $s \mapsto \mathcal{E}(\gamma(s))$ is λ'-convex, where $\lambda' = \lambda \mathsf{d}_{\mathcal{K}}(\gamma(0), \gamma(1))^2$. This property can be reformulated by local expressions using the second derivative in the form

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathcal{E}(\gamma) \geq \lambda\langle \mathcal{G}(\gamma)\gamma', \gamma'\rangle. \tag{3.7}$$

Using the first identity in (3.6) yields the identity

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathcal{E}(\gamma) = \frac{\mathrm{d}}{\mathrm{d}s}\langle \mathrm{D}\mathcal{E}(\gamma), \gamma'\rangle = \langle \gamma', \mathrm{D}^2\mathcal{E}(\gamma)\boldsymbol{\gamma}'\rangle + \Big\langle \mathrm{D}\mathcal{E}(\gamma), \frac{\mathrm{d}}{\mathrm{d}s}(\mathcal{K}(\gamma)\xi)\Big\rangle,$$

moreover, with the second identity in (3.6) we find

$$\langle \gamma', \mathrm{D}^2\mathcal{E}(\gamma)\gamma'\rangle + \langle \mathrm{D}\mathcal{E}(\gamma), \mathrm{D}\mathcal{K}(\gamma)[\gamma']\xi\rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}(\gamma)\left[\mathcal{K}(\gamma)\mathrm{D}\mathcal{E}(\gamma)\right]\xi\rangle \geq \lambda\langle \mathcal{G}(\gamma)\gamma', \gamma'\rangle.$$

From the definition of the vector field $u \mapsto \mathcal{F}(u) = \mathcal{K}(u)\mathrm{D}\mathcal{E}(u)$ we easily obtain

$$\langle \xi, \mathrm{D}\mathcal{F}(\gamma)v\rangle = \langle \mathrm{D}\mathcal{E}(\gamma), \mathrm{D}\mathcal{K}(\boldsymbol{\gamma})[v]\xi\rangle + \langle \xi, \mathrm{D}^2\mathcal{E}(\gamma)v\rangle.$$

Hence, using for $v = \gamma' = \mathcal{K}(\gamma)\xi$ we can rewrite (3.7) and we finally arrive at the estimate

$$\begin{aligned}&\langle \xi, \mathcal{M}(u)\xi\rangle \geq \lambda\langle \xi, \mathcal{K}(u)\xi\rangle \text{ for all } u \text{ and } \xi, \text{ where}\\ &\langle \xi, \mathcal{M}(u)\xi\rangle = \langle \xi, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\xi\rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi\rangle,\end{aligned}$$

which is the crucial estimate in (3.4).

Note that in the Wasserstein case $\mathcal{K}(u)\xi = -\mathrm{div}(u\nabla\xi)$ the operator $\mathcal{M}(u)$ is a fourth order differential operator (see examples in Section 3.3). Hence, to make the estimate in (3.4) well-defined we resort to dense subsets $\mathcal{Z} \subset X$.

3.2 Abstract setup

In this section we provide an abstract formulation such that the theory of [DaS08] can be applied to general systems $(X, \mathcal{E}, \mathcal{K})$, in particular to systems of partial differential equations, where $\mathcal{K}$ is allowed to be a partial differential operator as well. The main point of [DaS08] is that it is sufficient to establish the geodesic λ-convexity of $\mathcal{E}$ on a dense set, where all the calculations on functions can be done rigorously. Then, the abstract theory allows us to extend the geodesic λ-convexity of the functional $\mathcal{E}$ to the closure of the domain of $\mathcal{E}$.

We consider a set $\mathcal{X}$ which is a closed subset of a Banach space X, e.g. vectors of Radon measures. For the smooth solutions and their velocities we need smaller spaces

$$Z \subset Y \subset X$$

with dense and continuous embeddings. For $u \in Y$ the norm induced by the metric $\mathcal{G}(u)$ will be equivalent to that of a Hilbert space H, for which we assume

$$Y \subset H \text{ with dense and continuous embedding.}$$

We assume that open and connected sets $\mathcal{Z} \subset Z$ and $\mathcal{Y} \subset Y$ exist such that

$$\mathcal{Z} \subset Z \cap \mathcal{X}, \quad \mathcal{Z} \subset \mathcal{Y} \subset Y \cap \mathcal{X}, \text{ and } \mathcal{Z} \text{ is dense in } \mathcal{X}.$$

We refer to Section 3.3 for concrete examples of the various spaces.

We consider the gradient system restricted to the subset $\mathcal{Z}$, i.e., the triple $(\mathcal{Z}, \mathcal{E}, \mathcal{K})$ and assume that it satisfies

$$\mathcal{E} \in \mathrm{C}^2(\mathcal{Z}; \mathbb{R}), \quad \mathcal{K} \in \mathrm{C}^1\big(\mathcal{Y}; \mathrm{Lin}(H^*; H)\big), \quad \mathcal{G} = \mathcal{K}^{-1} \in \mathrm{C}^1\big(\mathcal{Y}; \mathrm{Lin}(H; H^*)\big), \tag{3.8}$$

where we additionally assume that $\mathcal{E}$ is bounded from below.

Thus, the evolution of the system reads

$$\dot{u} = -\mathcal{F}(u) \overset{\text{def}}{=} -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u),$$

where, having in mind PDEs, we assume the smoothness of the vector field $\mathcal{F}$

$$\mathcal{F} \in \mathrm{C}^1(\mathcal{Z}; Y) \text{ and } \mathrm{D}\mathcal{F} \in \mathrm{C}^0\big(\mathcal{Z}; \mathrm{Lin}(Z; Y)\big) \cap \mathrm{C}^0\big(\mathcal{Z}; \mathrm{Lin}(Y; H)\big), \tag{3.9}$$

which is what one would obtain composing the smoothness of $\mathcal{K}$ and $\mathcal{E}$ in (3.8). In particular, with the assumptions above the quadratic form $\langle \xi, \mathcal{M}(u)\xi \rangle$ is well-defined for $u \in \mathcal{Z}$ and $\xi \in \mathcal{G}(u)Y = \{\eta \in H^* : \mathcal{K}(u)\eta \in Y\}$.

3.2.1 Geodesic curves and geodesic $\boldsymbol{\lambda}$-convexity

The metric tensor $\mathcal{G} = \mathcal{K}^{-1}$ generates a distance $\mathsf{d}_{\mathcal{K}} : X \times X \to [0, \infty]$ in the usual way: For $u_0, u_1 \in X$ we define the set of connecting curves via

$$\boldsymbol{C}(u_0, u_1) = \Big\{ \gamma \in \mathrm{C}^1([0,1]; X) : \gamma(0) = u_0, \ \gamma(1) = u_1 \Big\}.$$

This allows us to define the distance $\mathsf{d}_{\mathcal{K}}$ as follows

$$\begin{aligned} &\mathsf{d}_{\mathcal{K}}(u_0, u_1)^2 = \inf \{ J_{\mathcal{K}}(\gamma) : \gamma \in \boldsymbol{C}(u_0, u_1) \} \\ &\text{with} \quad J_{\mathcal{K}}(\gamma) = \int_0^1 \mathcal{A}(\gamma(s), \gamma'(s)) \, \mathrm{d}s. \end{aligned} \tag{3.10}$$

Here, γ' denotes the derivative with respect to the arclength parameter s, and $\mathcal{A}$ is the action functional given by

$$\mathcal{A}(u, v) = \begin{cases} \langle \mathcal{G}(u)v, v \rangle & \text{if } (u, v) \in \mathcal{Y} \times H, \\ +\infty & \text{else.} \end{cases}$$

It is easy to see that $\mathsf{d}_\mathcal{K}$ is symmetric and satisfies the triangle inequality. We assume positivity, i.e.,

$$\forall\, u, w \in \mathcal{Z}: \quad u \neq w \implies \mathsf{d}_\mathcal{K}(u,w) > 0. \tag{3.11}$$

Thus, we may consider also the metric gradient system $(X, \mathcal{E}, \mathsf{d}_\mathcal{K})$ in the sense of [AGS05]. We refer to the latter or to [CL*10] for distances $\mathsf{d}_\mathcal{K}$ in more general cases. As in any metric space $(X, \mathsf{d}_\mathcal{K})$, a geodesic curve connecting u_0 and u_1 is a curve $\gamma \in \boldsymbol{C}(u_0,u_1)$ satisfying

$$\forall\, r,s \in [0,1]: \quad \mathsf{d}_\mathcal{K}(\gamma(r),\gamma(s)) = |r{-}s|\, \mathsf{d}_\mathcal{K}(u_0,u_1). \tag{3.12}$$

Remark 3.2.1 *If $\mathcal{Y}$ is a convex subset of $Y \subset X$ and $\mathcal{Y} \ni u \mapsto \langle \eta, \mathcal{K}(u)\eta\rangle$ is concave for all η, then $(u,v) \mapsto \langle \mathcal{G}(u)v, v\rangle$ is (jointly) convex on $\mathcal{Y} \times H$. As a consequence the functional $J_\mathcal{K}$ in* (3.10) *and hence $\mathsf{d}_\mathcal{K}^2 : \mathcal{Y} \times \mathcal{Y} \to [0,\infty[$ is convex as well.*

For a given $\lambda \in \mathbb{R}$, a functional $\mathcal{E}$ is called *geodesically λ-convex* with respect to the metric $\mathsf{d}_\mathcal{K}$ if for all geodesics $\gamma : [s_a, s_b] \to X$ we have

$$\mathcal{E}(\gamma(s_\theta)) \leq (1{-}\theta)\mathcal{E}(\gamma(s_0)) + \theta\mathcal{E}(\gamma(s_1)) - \lambda\frac{\theta(1{-}\theta)}{2}\mathsf{d}_\mathcal{K}(\gamma(s_0),\gamma(s_1))^2 \tag{3.13}$$

for all $\theta \in [0,1]$ and $s_0, s_1 \in [s_a, s_b]$, where $s_\theta = (1{-}\theta)s_0 + \theta s_1$.

As we have seen in the previous chapter it is most natural to model reaction-diffusion systems in terms of the Onsager operator. Hence, we will formulate the convexity conditions in terms of $\mathcal{E}$, $\mathcal{K}$, and the vector field $\mathcal{F}$. However, from the mathematical point of view the metric $\mathcal{G} = \mathcal{K}^{-1}$ and the induced distance $\mathsf{d}_\mathcal{K}$ are important as well. Following the famous Benamou-Brenier formulation [BeB00] we can characterize our $\mathcal{G}$ in a similar fashion

$$\begin{aligned}\langle \mathcal{G}(\boldsymbol{u})\boldsymbol{v}, \boldsymbol{v}\rangle = \inf\Big\{ \int_\Omega \boldsymbol{\Sigma}{:}\mathbb{M}(\boldsymbol{u})^{-1}\boldsymbol{\Sigma} + \boldsymbol{\sigma}\cdot\mathbb{K}(\boldsymbol{u})^{-1}\boldsymbol{\sigma}\,\mathrm{d}x \;\Big|\; &\boldsymbol{\Sigma} \in \mathrm{L}^2(\Omega;\mathbb{R}^{I\times d}),\\ &\boldsymbol{\sigma} \in \mathrm{L}^2(\Omega;\mathbb{R}^I),\ \boldsymbol{\sigma} - \mathrm{div}\boldsymbol{\Sigma} = \boldsymbol{v}\Big\}.\end{aligned} \tag{3.14}$$

In particular, concavity of the tensors $\mathbb{M}$ and $\mathbb{K}$ (i.e. for all $\boldsymbol{\xi}$ the mapping $\boldsymbol{u} \mapsto \boldsymbol{\xi}\cdot\mathbb{K}(\boldsymbol{u})\boldsymbol{\xi}$ is concave) we find that $(\boldsymbol{u},\boldsymbol{v}) \mapsto \langle \mathcal{G}(\boldsymbol{u})\boldsymbol{v}, \boldsymbol{v}\rangle$ is convex, which can be used to establish the existence of geodesic curves.

3.2.2 A simple example

Only in very few cases $\mathsf{d}_\mathcal{K}$ can be calculated explicitly, all relying on the Wasserstein distance d_Wass, see [AGS05, Vil09]. For constants $\mu \geq 0$ and $\kappa \geq 0$ consider the Onsager operator $\mathcal{K}_{\mu,\kappa}(u)\xi = -\mathrm{div}(\mu u \nabla\xi) + \kappa u\xi$, which is affine in u. The case $\kappa = 0$ corresponds to the Wasserstein distance, i.e., we have on the set $X = \{u \in \mathsf{Meas}(\Omega) : u \geq 0\}$ of nonnegative Radon measures the distance

$$\mathsf{d}_{\mathcal{K}_{\mu,0}}(u_0,u_1) = \begin{cases} \sqrt{\alpha/\mu}\;\mathsf{d}_\mathrm{Wass}(u_0/\alpha, u_1/\alpha) & \text{if } \mathrm{vol}(u_0) = \mathrm{vol}(u_1) = \alpha, \\ +\infty & \text{otherwise.} \end{cases}$$

For $\kappa = 0$ the Onsager operator $\mathcal{K}_{\mu,0}$ is mass preserving, hence X decomposes into the components $X_\alpha = \{u \in X : \mathrm{vol}(u) = \alpha\}$. For $\mu = 0$ there is no spatial interaction, and we find the explicit formula

$$\mathsf{d}_{\mathcal{K}_{0,\kappa}}(u_0, u_1) = \sqrt{\tfrac{4}{\kappa}}\, \|\sqrt{u_0} - \sqrt{u_1}\|_{\mathrm{L}^2(\Omega)}.$$

This distance is related to the Kakutani-Hellinger distance of order 1/2 on the space of probability measures (see [Hel09, Kak48]), where it induces the same topology as the total variation. For a survey on this distance we refer to [LiV87].

Arguing as in [BeB00] we introduce a space-time dependent Lagrange multiplier $\eta(t, x)$ for the constraints in (3.14) to obtain after integration by parts

$$\mathsf{d}_{\mathcal{K}_{\mu,\kappa}}(u_0, u_1)^2 = \inf_{u,\Sigma,\sigma} \sup_{\eta} \Big\{ \int_0^1 \int_\Omega \left[\frac{|\Sigma|^2}{2\mu u} + \frac{|\sigma|^2}{2\kappa u} - \dot\eta u - \nabla\eta \cdot \Sigma - \sigma\eta \right] \mathrm{d}x\,\mathrm{d}t \\ - \int_\Omega [\eta(0, x)u_0 - \eta(1, x)u_1]\,\mathrm{d}x \Big\}. \quad (3.15)$$

Now, we observe that for positive u we have pointwise in time and space

$$\frac{|\Sigma|^2}{2\mu u} + \frac{|\sigma|^2}{2\kappa u} = \sup_{C_{\mu,\kappa}} \{au + b \cdot \Sigma + c\sigma\},$$

where $C_{\mu,\kappa} = \Big\{(a, b, c) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} : a + \frac{\mu}{2}|b|^2 + \frac{\kappa}{2}|c|^2 \le 0\Big\}$ is convex. Hence, we can rewrite (3.15) as

$$\mathsf{d}_{\mathcal{K}_{\mu,\kappa}}(u_0, u_1)^2 = \inf_{u,\Sigma,\sigma} \sup_{\eta} \sup_{a,b,c} \Big\{ I_{C_{\mu,\kappa}}(a, b, c) + \int_0^1 \int_\Omega u(a - \dot\eta) + \Sigma \cdot (b - \nabla\eta) + \sigma(c - \eta)\,\mathrm{d}x\,\mathrm{d}t \\ \int_\Omega [\eta(1, x)u_1 - \eta(0, x)u_0]\,\mathrm{d}x \Big\},$$

where $I_{C_{\mu,\kappa}}$ is the indicator function of the set $C_{\mu,\kappa}$. Hence, assuming that we are allowed to interchange inf and sup in the above equation we conjecture the formula

$$\mathsf{d}_{\mathcal{K}_{\mu,\kappa}}(u_0, u_1)^2 = \sup \left\{ \textstyle\int_\Omega \eta(1, x)u_1(\mathrm{d}x) - \int_\Omega \eta(0, x)u_0(\mathrm{d}x) : \dot\eta + \frac{\mu}{2}|\nabla\eta|^2 + \frac{\kappa}{2}\eta^2 \le 0 \right\}.$$

This and other characterizations of reaction-diffusion distances will be investigated in subsequent work.

3.2.3 Properties of geodesically λ-convex gradient flows

In this section we collect some useful properties of geodesically λ-convex systems. We refer to [DaS08] for the full discussion. First, we have a Lipschitz continuous dependence of the solutions u_j, $j = 1, 2$, on the initial data, namely

$$\text{for all } t \geq 0 : \qquad \mathsf{d}_{\mathcal{K}}(u_1(t), u_2(t)) \leq \mathrm{e}^{-\lambda t}\mathsf{d}_{\mathcal{K}}(u_1(0), u_2(0)). \tag{3.16}$$

In particular, for $\lambda \geq 0$ we have a contraction semigroup. If $\lambda > 0$ we obtain exponential decay towards the unique equilibrium state u_*, which minimizes $\mathcal{E}$, i.e.,

$$\mathsf{d}_{\mathcal{K}}(u(t), u_*) \leq \mathrm{e}^{-\lambda t}\mathsf{d}_{\mathcal{K}}(u(0), u_*).$$

It was shown in [DaS08, Prop. 3.1] that for geodesically λ-convex functionals the solutions of the (differential) gradient flow (2.1) satisfy a purely metric formulation in terms of the evolutionary variational inequality (EVI$_\lambda$)

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathsf{d}_{\mathcal{K}}^2(u(t), w) + \frac{\lambda}{2}\mathsf{d}_{\mathcal{K}}^2(u(t), w) + \mathcal{E}(u(t)) \leq \mathcal{E}(w), \quad \forall\, w \in X,\ t > 0,$$

where for a function $f : [0, \infty[\ \to \mathbb{R}$ we set $\frac{\mathrm{d}^+}{\mathrm{d}t}f(t) = \limsup_{h\to 0^+} \frac{1}{h}\big(f(t{+}h) - f(t)\big)$. The above differential form is (for $t \mapsto \mathcal{E}(u(t))$ decreasing) equivalent to the integrated form of (EVI$_\lambda$) given by

$$\frac{\mathrm{e}^{\lambda\tau}}{2}\mathsf{d}_{\mathcal{K}}(u(t{+}\tau), w)^2 - \frac{1}{2}\mathsf{d}_{\mathcal{K}}(u(t), w)^2 \leq \mathsf{E}_\lambda(\tau)\big(\mathcal{E}(w) - \mathcal{E}(u(t{+}\tau))\big) \quad \forall\, w \in X,\ t, \tau \geq 0,$$

where $\mathsf{E}_\lambda(t) = \int_0^t \exp(\lambda r)\,\mathrm{d}r$ (see [DaS08, Prop. 3.1]). In particular, the solutions of (EVI$_\lambda$) satisfy for $\lambda \neq 0$ the uniform regularization bound

$$\mathcal{E}(u(t)) \leq \mathcal{E}(w) + \frac{1}{2\mathsf{E}_\lambda(t)}\mathsf{d}_{\mathcal{K}}(u(0), w)^2 \qquad \forall\, w \in X,\ t > 0.$$

Moreover, the solutions are uniformly continuous in time:

$$\mathsf{d}_{\mathcal{K}}(u(t{+}\tau), u(t))^2 \leq 2\mathsf{E}_{-\lambda}(\tau)\Big(\mathcal{E}(u(t)) - \inf_{w\in X} \mathcal{E}(w)\Big).$$

One of the main observations in [DaS08] is that the existence of a flow map $t \mapsto \mathcal{S}_t(u) = u(t)$ satisfying (EVI$_\lambda$) yields the geodesic λ-convexity of the functional $\mathcal{E}$. More precisely, we have the following (see [DaS08, Theorem 3.2]).

Theorem 3.2.2 (Daneri & Savaré [DaS08]) *Assume that $t \mapsto \mathcal{S}_t(u) = u(t)$ solves (EVI$_\lambda$) and $t \mapsto \mathcal{E}(u(t))$ is decreasing. If $\gamma : [0, 1] \to X$ is a Lipschitz curve connecting $u_0, u_1 \in X$ and satisfying for $0 \leq r, s \leq 1$ the estimate*

$$\mathsf{d}_{\mathcal{K}}(\gamma(r), \gamma(s)) \leq L|r - s|, \qquad \text{with } L^2 \leq \mathsf{d}_{\mathcal{K}}(u_0, u_1)^2 + \varepsilon^2 \tag{3.17}$$

for some constant $\varepsilon \geq 0$, then for every $t \geq 0$ and $s \in [0, 1]$

$$\mathcal{E}\big(\mathcal{S}_t(\gamma(s))\big) \leq (1{-}s)\mathcal{E}(u_0) + s\mathcal{E}(u_1) - \frac{\lambda s(1{-}s)}{2}\mathsf{d}_{\mathcal{K}}(u_0, u_1)^2 + \frac{\varepsilon^2}{2\mathsf{E}_\lambda(t)}s(1{-}s).$$

In particular, when γ is a geodesic curve (i.e. $L = \mathsf{d}_{\mathcal{K}}(u_0,u_1)$ and $\varepsilon = 0$ in (3.17)*), then $\mathcal{E}$ satisfies*

$$\mathcal{E}(\gamma(s)) \leq (1{-}s)\mathcal{E}(u_0) + s\mathcal{E}(u_1) - \frac{\lambda s(1{-}s)}{2}\mathsf{d}_{\mathcal{K}}(u_0,u_1)^2;$$

thus, $\mathcal{E}$ is geodesically λ-convex.

3.2.4 Completion of smooth gradient flows

In addition to the assumptions in (3.8) and (3.9) we now assume that the triple $(\mathcal{Z},\mathcal{E},\mathcal{K})$ generates a global semiflow in $\mathcal{Z}$ in the form $u(t) = \mathcal{S}_t(u(0))$ for $t > 0$ with a semigroup $\mathcal{S} : [0,\infty[\, \times \mathcal{Z} \to \mathcal{Z}$, i.e.,

$$\begin{aligned}&\mathcal{S}_t \circ \mathcal{S}_r = \mathcal{S}_{t+r} \text{ for } r,t \geq 0;\\ &\mathcal{S}_t(u) \to u \text{ in } Z \quad\text{and}\quad \tfrac{1}{t}(\mathcal{S}_t(u){-}u) \to -\mathcal{F}(u) \text{ in } Y \qquad \text{for } t \to 0^+.\end{aligned}$$

More precisely, we make the following regularity assumptions on the semigroup $\mathcal{S}$

$$\mathcal{S} \in \mathrm{C}^0([0,\infty[\,{\times}\mathcal{Z};Z) \,\cap\, \mathrm{C}^1([0,\infty[\,{\times}\mathcal{Z};Y) \,\cap\, \mathrm{C}^2([0,\infty[\,{\times}\mathcal{Z};H). \tag{3.18}$$

In particular, this implies that $\mathrm{D}\mathcal{S}$ and $\mathcal{F}(u) = -\partial_t\mathcal{S}_t(u)|_{t=0}$ satisfy

$$(t,u) \mapsto \mathrm{D}\mathcal{S}_t(u) \in \mathrm{C}^0([0,\infty[\,{\times}\mathcal{Z};\mathrm{Lin}(Z;Y)) \,\cap\, \mathrm{C}^1([0,\infty[\,{\times}Z;\mathrm{Lin}(Z;H)). \tag{3.19}$$

We define the functionals $\mathcal{A} : \mathcal{Y} \times H \to \mathbb{R}$ and $\mathcal{B} : \mathcal{Z} \times Y \to \mathbb{R}$ via

$$\mathcal{A}(u,v) = \langle \mathcal{G}(u)v, v\rangle, \quad \mathcal{B}(u,v) = \langle \mathcal{G}(u)v, \mathrm{D}\mathcal{F}(u)v\rangle + \frac{1}{2}\langle \mathrm{D}\mathcal{G}(u)[\mathcal{F}(u)]v, v\rangle$$

and obtain the following formulas.

Proposition 3.2.3 *(i) For $u \in \mathrm{C}^1([t_0,t_1];\mathcal{Y})$ and $v \in \mathrm{C}^1([t_0,t_1];H)$ we have*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(u(t),v(t)) = 2\langle \mathcal{G}(u)v, \dot{v}\rangle + \langle \mathrm{D}\mathcal{G}(u)[\dot{u}]v, v\rangle. \tag{3.20}$$

(ii) For all $u \in \mathcal{Z}$, $v \in Z$, and $t \geq 0$ we have

$$\tfrac{1}{2}\,\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{A}(\mathcal{S}_t(u),\mathrm{D}\mathcal{S}_t(u)v) + \mathcal{B}(\mathcal{S}_t(u),\mathrm{D}\mathcal{S}_t(u)v) = 0. \tag{3.21}$$

Proof: Part (i) follows simply by the assumed smoothness of $\mathcal{G}$ and the chain rule for the Fréchet derivative in Banach spaces. Part (ii) is an application of part (i) by using $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}_t(u) = -\mathcal{F}(\mathcal{S}_t(u))$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{D}\mathcal{S}_t(u) = -\mathrm{D}\mathcal{F}(\mathcal{S}_t(u))\mathrm{D}\mathcal{S}_t(u)$. $\square$

The central idea of [DaS08] is the transport of curves $\gamma_t \in \boldsymbol{C}(u_0,\mathcal{S}_t(u_1))$ defined via

$$\gamma_t(s) = \mathcal{S}_{st}(\gamma(s)) \quad \text{for } \gamma \in \boldsymbol{C}(u_0,u_1) \cap \mathrm{C}^2([0,1];\mathcal{Z}).$$

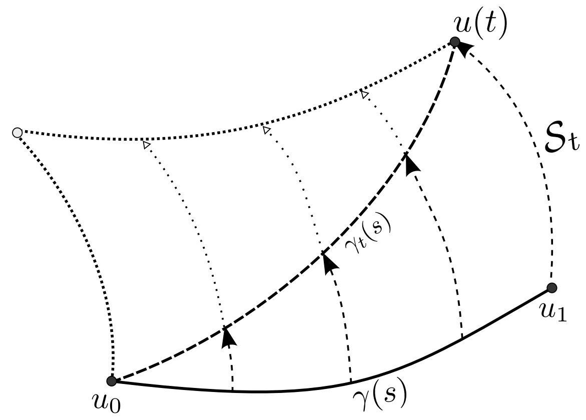

Figure 3.1: Variation of the curve $s \mapsto \gamma(s)$ under the semigroup $\mathcal{S}_t$

Note, in particular, that the endpoint γ_0 remains fixed, i.e., $\gamma_0(s) \equiv \gamma_0$. The main tool is the following relation (3.22) for the functions

$$A(s,t) \overset{\text{def}}{=} \mathcal{A}(\gamma_t(s), \gamma_t'(s)), \quad B(s,t) \overset{\text{def}}{=} \mathcal{B}(\gamma_t(s), \gamma_t'(s)), \quad \text{and } E(s,t) \overset{\text{def}}{=} \mathcal{E}(\gamma_t(s)),$$

where $\gamma_t'(s) = \partial_s(\gamma_t(s)) \in Y$ denotes the derivative with respect to the arclength parameter s.

Proposition 3.2.4 *For every curve $\gamma \in \boldsymbol{C}(w,u)$ we have*

$$\frac{1}{2}\frac{\partial}{\partial t}A(s,t) + \frac{\partial}{\partial s}E(s,t) + sB(s,t) = 0. \tag{3.22}$$

Proof: We first observe that the mapping $\Gamma : (s,t) \mapsto \gamma_t(s)$ satisfies

$$\Gamma \in \mathrm{C}^0([0,1]\times [0,\infty[\,;\mathcal{Z}) \cap \mathrm{C}^1([0,1]\times [0,\infty[\,;Y) \cap \mathrm{C}^2([0,1]\times [0,\infty[\,;H).$$

In particular, using the definition of the semiflow $\mathcal{S}_t$ we have the relations

$$\partial_t\gamma_t(s) = -s\mathcal{F}(\gamma_t(s)) \text{ and } \partial_t(\gamma_t'(s)) = \partial_s\partial_t\gamma_t(s) = -\mathcal{F}(\gamma_t(s)) - s\mathrm{D}\mathcal{F}(\gamma_t(s))\gamma_t'(s).$$

Note that we will not need an expression for $\gamma_t'(s)$. Applying Proposition 3.2.3(i) and the above formulas for $\partial_t\gamma_t(s)$ and $\partial_t(\gamma_t'(s))$ we find

$$\begin{aligned}
\tfrac{1}{2}\tfrac{\partial}{\partial t}A(s,t) &= -\big\langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), \mathcal{F}(\gamma_t(s))\big\rangle - \big\langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), s\,\mathrm{D}\mathcal{F}(\gamma_t(s))\gamma_t'(s)\big\rangle \\
&\quad - \tfrac{1}{2}\big\langle \mathrm{D}\mathcal{G}(\gamma_t(s))[s\mathcal{F}(\gamma_t(s))]\gamma_t'(s), \gamma_t'(s)\big\rangle \\
&= -\big\langle \mathcal{G}(\gamma_t(s))\gamma_t'(s), \mathcal{K}(\gamma_t(s))\mathrm{D}\mathcal{E}(\gamma_t(s))\big\rangle - s\mathcal{B}(\gamma_t(s), \gamma_t'(s)) \\
&= -\big\langle \mathrm{D}\mathcal{E}(\gamma_t(s)), \gamma_t'(s)\big\rangle - sB(s,t) \;=\; -\tfrac{\partial}{\partial s}E(s,t) - sB(s,t),
\end{aligned}$$

which is the desired result. $\square$

One of the main achievements of [DaS08] was to show that the identity (3.22) can be used to derive the evolutionary variational inequality (EVI_λ), namely

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathsf{d}_{\mathcal{K}}(\mathcal{S}_t(u),w)^2 + \frac{\lambda}{2}\mathsf{d}_{\mathcal{K}}(\mathcal{S}_t(u),w)^2 + \mathcal{E}(\mathcal{S}_t(u)) \le \mathcal{E}(w) \quad \forall\, w\in\mathcal{Z},\ t\ge 0. \tag{3.23}$$

It is especially interesting that this result holds without any completeness of the space $\mathcal{Z}$. The crucial assumption needed is that $B(s,t)$ can be estimated in terms of $A(s,t)$, namely in the form $B(s,t)\ge\lambda A(s,t)$ along the curves γ_t. The following result is an abstract version of the ideas in [DaS08].

Theorem 3.2.5 *Assume that* $(\mathcal{Z},\mathcal{E},\mathcal{K})$ *generates the semigroup* $\mathcal{S}$ *and the above conditions* (3.8)—(3.18) *hold. If additionally*

$$\begin{aligned}&\forall\, u\in\mathcal{Z}\ \forall\, v\in Y: \quad \mathcal{B}(u,v)\ge\lambda\mathcal{A}(u,v),\\ &\textit{i.e.}\ \ \langle\mathcal{G}(u)v,\mathrm{D}\mathcal{F}(u)v\rangle+\frac{1}{2}\langle\mathrm{D}\mathcal{G}(u)[\mathcal{F}(u)]v,v\rangle\ge\lambda\langle\mathcal{G}(u)v,v\rangle,\end{aligned} \tag{3.24}$$

then, the semigroup $\mathcal{S}$ *satisfies* (EVI_λ) *given in* (3.23).

Proof: We follow the steps in the proof of [DaS08, Theorem 5.1], where the underlying metric space (X,d) is *not* assumed to be complete. Hence, we are able to choose the smaller metric space $(\mathcal{Z},\mathsf{d}_{\mathcal{K}})$.

For $w,u_0\in\mathcal{Z}$ let $\gamma\in\boldsymbol{C}(w,u_0)\cap\mathrm{C}^2([0,1];\mathcal{Z})$ be given and define the family of curves $s\mapsto\gamma_t(s)=\mathcal{S}_{st}(\gamma(s))$ and $u(t)=\gamma_t(1)=\mathcal{S}_t(u_1)$ as above. The identity in Proposition 3.2.4 and estimate (3.24) for $u=\gamma_t(s)\in\mathcal{Z}$ and $v=\gamma_t'(s)\in Y$ yield the estimate

$$\frac{1}{2}\frac{\partial}{\partial t}A(s,t)+\lambda sA(s,t)+\frac{\partial}{\partial s}E(s,t)\le 0.$$

Multiplying this estimate by $(s,t)\mapsto\exp(2\lambda st)>0$ we obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\Big(\mathrm{e}^{2\lambda st}A(s,t)\Big)+\frac{\partial}{\partial s}\Big(\mathrm{e}^{2\lambda st}E(s,t)\Big)\le 2\lambda t\mathrm{e}^{2\lambda st}E(s,t).$$

We define the auxiliary function $\mathsf{E}_{2\lambda}:[0,\infty[\,\to[0,\infty[$ by

$$\mathsf{E}_{2\lambda}(t)=\int_0^t\mathrm{e}^{2\lambda r}\,\mathrm{d}r=\begin{cases}\frac{\mathrm{e}^{2\lambda t}-1}{2\lambda} & \text{if } \lambda\ne 0,\\ t & \text{if } \lambda=0.\end{cases}$$

Integrating the estimate above with respect to s over $[0,1]$ and a further integration with respect to t over $[0,\tau]$ gives

$$\begin{aligned}\frac{1}{2}\int_0^1\mathrm{e}^{2\lambda s\tau}A(s,\tau)\,\mathrm{d}s-\frac{1}{2}\int_0^1\mathcal{A}(\gamma(s),\gamma'(s))\,\mathrm{d}s+\mathsf{E}_{2\lambda}(\tau)\mathcal{E}(u(\tau))&\\ \le\tau\mathcal{E}(w)+\int_0^1\int_0^\tau 2\lambda t\mathrm{e}^{2\lambda st}E(s,t)\,\mathrm{d}t\,\mathrm{d}s,&\end{aligned} \tag{3.25}$$

where we have used the identities $E(0,t)=\mathcal{E}(\gamma_t(0))=\mathcal{E}(w)$, $A(s,0)=\mathcal{A}(\gamma(s),\gamma'(s))$ and that the map $t\mapsto E(1,t)=\mathcal{E}(u(t))$ is decreasing, in particular $\mathcal{E}(u(t))\geq\mathcal{E}(u(\tau))$ for $0\leq t\leq\tau$.

Part (i) of Lemma 3.2.6 below for $f_t(s)=\exp(2\lambda st)$ yields the estimate

$$\frac{\sigma_\lambda(\tau)\mathrm{e}^{\lambda\tau}}{2}\mathsf{d}_{\mathcal{K}}(u(\tau),w)^2\leq\frac{1}{2}\int_0^1\mathrm{e}^{2\lambda s\tau}A(s,\tau)\,\mathrm{d}s, \tag{3.26}$$

where $\sigma_\lambda(t)=\lambda t/\sinh(\lambda t)$. Moreover, we can assume that for a fixed $\varepsilon>0$ the curve $\gamma=\gamma_\varepsilon$ satisfies the estimate

$$\int_0^1\mathcal{A}(\gamma_\varepsilon(s),\gamma_\varepsilon'(s))\,\mathrm{d}s\leq\mathsf{d}_{\mathcal{K}}(w,u_0)^2+\varepsilon^2. \tag{3.27}$$

Moreover, by a standard reparametrization technique (see next Lemma 3.2.6), we can also assume that γ_ε is almost a constant speed geodesic, namely

$$\forall r,s\in[0,1]:\quad\mathsf{d}_{\mathcal{K}}(\gamma_\varepsilon(r),\gamma_\varepsilon(s))\leq L_\varepsilon|r-s|,\qquad L_\varepsilon^2\overset{\text{def}}{=}\mathsf{d}_{\mathcal{K}}(w,u_0)^2+\varepsilon^2. \tag{3.28}$$

Using the estimates (3.26),(3.27) in (3.25) we obtain

$$\begin{aligned}\frac{\sigma_\lambda(\tau)\mathrm{e}^{\lambda\tau}}{2}\mathsf{d}_{\mathcal{K}}(u(\tau),w)^2-\frac{1}{2}\mathsf{d}_{\mathcal{K}}(u_0,w)^2+\mathsf{E}_{2\lambda}(\tau)\mathcal{E}(u(\tau))&\\ \leq\tau\mathcal{E}(w)+\int_0^1\int_0^\tau 2\lambda t\mathrm{e}^{2\lambda st}E(s,t)\,\mathrm{d}t\,\mathrm{d}s+\frac{\varepsilon^2}{2}.&\end{aligned} \tag{3.29}$$

Now, the cases $\lambda\leq0$ and $\lambda>0$ have to be treated differently.

<u>1. Let us first consider the case $\lambda\leq0$:</u> Since $\mathcal{E}$ is assumed to be bounded from below, say by a constant $C_{\mathcal{E}}$, we have

$$\int_0^1\int_0^\tau 2\lambda t\mathrm{e}^{2\lambda st}E(s,t)\,\mathrm{d}t\,\mathrm{d}s\leq 2C_{\mathcal{E}}\lambda\int_0^1\int_0^\tau t\mathrm{e}^{2\lambda st}\,\mathrm{d}t\,\mathrm{d}s=C_{\mathcal{E}}\Big(\frac{\mathrm{e}^{2\lambda\tau}-1}{2\lambda}-\tau\Big)=:C_{\mathcal{E}}\mathsf{F}_\lambda(\tau).$$

With this and since ε is arbitrary we divide by τ and use $\frac{1}{\tau}\mathsf{F}_\lambda(\tau)\to0$ and $\frac{1}{\tau}\mathsf{E}_{2\lambda}(\tau)\to1$ for $\tau\downarrow0$ to arrive at

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}\tau}\Big(\sigma_\lambda(\tau)\mathrm{e}^{\lambda\tau}\mathsf{d}_{\mathcal{K}}(u(\tau),w)^2\Big)\Big|_{\tau=0}+\mathcal{E}(u_0)\leq\mathcal{E}(w).$$

Since $\sigma_\lambda'(0)=0$ it is then easy to check that

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}\tau}\Big(\sigma_\lambda(\tau)\mathrm{e}^{\lambda\tau}\mathsf{d}_{\mathcal{K}}(u(\tau),w)^2\Big)\Big|_{\tau=0}=\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}\tau}\mathsf{d}_{\mathcal{K}}(u(\tau),w)\big|_{\tau=0}+\frac{\lambda}{2}\mathsf{d}_{\mathcal{K}}(u_0,w),$$

which yields (EVI_λ) for $\tau=0$. For positive τ the result follows from the semigroup property of $\mathcal{S}$

2. Let us now consider the case $\lambda > 0$: Note that if (3.24) holds for $\lambda > 0$ then it obviously also holds for $\lambda = 0$ hence we can argue as in the first step to obtain (EVI_0). Due to (3.28) we can apply Theorem 3.2.2 to obtain

$$\begin{aligned} tE(s,t) = t\mathcal{E}(\gamma_t(s)) &\le t\Big((1-s)\mathcal{E}(w) + s\mathcal{E}(u_0) + \frac{\varepsilon^2 s(1-s)}{2t}\Big) \\ &\le t(\mathcal{E}(w)+\mathcal{E}(u_0)) + \frac{\varepsilon^2}{2} \end{aligned}$$

since $s \in [0,1]$. Thus, we get

$$\int_0^\tau \int_0^1 2\lambda t \mathrm{e}^{2\lambda s t} E(s,t)\,\mathrm{d}s\,\mathrm{d}t \le \lambda\tau\mathrm{e}^{2\lambda\tau}\Big(\tau(\mathcal{E}(u_1)+\mathcal{E}(u_0)) + \varepsilon^2\Big),$$

where we used that $\mathrm{e}^{2\lambda st} \le \mathrm{e}^{2\lambda\tau}$ for $0 \le s \le 1$ and $0 \le t \le \tau$. Inserting this estimate in (3.29) and letting $\varepsilon \downarrow 0$ we find

$$\frac{1}{2\tau}\Big(\sigma_\lambda(\tau)\mathrm{e}^{\lambda\tau}\mathsf{d}_\mathcal{K}(u(\tau),w)^2 - \mathsf{d}_\mathcal{K}(u_0,w)^2\Big) + \frac{\mathsf{E}_{2\lambda}(\tau)}{\tau}\mathcal{E}(u(\tau)) \le \lambda\tau\mathrm{e}^{2\lambda\tau}\big(\mathcal{E}(u_0)+\mathcal{E}(w)\big).$$

Letting $\tau \downarrow 0$, the term in the right-hand side vanishes, such that we obtain the (EVI_λ) also in the case in which $\lambda > 0$. □

The following reparametrization lemma, which was used in the proof of Theorem 3.2.5, is a generalized version of Lemma 5.1 in [DaS08]. For the convenience of the reader we provide the proof here.

Lemma 3.2.6 ([DaS08, Lemma 5.1]) *For $u, w \in \mathcal{Z}$ let $\gamma \in \boldsymbol{C}(u,w) \cap \mathrm{C}^2([0,1];\mathcal{Z})$.*

(i) For every positive function $f \in \mathrm{C}^1([0,1])$ it holds that

$$\mathsf{d}_\mathcal{K}(u,w)^2 \le M_f \int_0^1 f(s)\mathcal{A}(\gamma(s),\gamma'(s))\,\mathrm{d}s, \quad \textit{where} \quad M_f \overset{def}{=} \int_0^1 \frac{1}{f(s)}\,\mathrm{d}s. \tag{3.30}$$

(ii) Moreover, for every $\varepsilon > 0$ there exists a smooth rescaling $\kappa_\varepsilon : [0,1] \to [0,1]$ so that the reparametrized families $\gamma_\varepsilon = \gamma \circ \kappa_\varepsilon$ satisfy $\gamma_\varepsilon \in \boldsymbol{C}(u,w) \cap \mathrm{C}^2([0,1],\mathcal{Z})$ and

$$\mathsf{d}_\mathcal{K}(\gamma_\varepsilon(s_0),\gamma_\varepsilon(s_1)) \le L|s_0-s_1|, \quad \textit{with} \quad L^2 \le \int_0^1 \mathcal{A}(\gamma(s),\gamma'(s))\,\mathrm{d}s + \varepsilon^2. \tag{3.31}$$

Proof: We consider the smooth and increasing map $\theta : [0,1] \to [0,1]$ given by $\theta(s) = \frac{1}{M_f}\int_0^s 1/f(r)\,\mathrm{d}r$. Moreover, let us denote by $\kappa = \theta^{-1}$ its inverse such that $\kappa'(\theta(s)) = M_f f(s)$. Then, we check that for the reparametrized curve $\overline{\gamma} \in \boldsymbol{C}(u,w) \cap \mathrm{C}^2([0,1];\mathcal{Z})$ given by $\overline{\gamma}(s) = \gamma(\kappa(s))$ it follows that

$$\mathsf{d}_\mathcal{K}(u,w)^2 \le \int_0^1 \mathcal{A}(\overline{\gamma}(r),\overline{\gamma}'(r))\,\mathrm{d}r = M_f \int_0^1 f(s)\mathcal{A}(\gamma(s),\gamma'(s))\,\mathrm{d}s,$$

which proves (3.30). Next, for $a(s) = \mathcal{A}(\gamma(s), \gamma'(s))$ we define the family $f_\varepsilon : [0,1] \to \mathbb{R}$ by

$$f_\varepsilon(s) = \frac{1}{\sqrt{\varepsilon^2 + a(s)}}, \quad \text{such that} \quad M_{f_\varepsilon} = \int_0^1 \sqrt{\varepsilon^2 + a(s)}\,\mathrm{d}s, \quad M_{f_\varepsilon}^2 \le \varepsilon^2 + \int_0^1 a(s)\,\mathrm{d}s.$$

Hence, we have that

$$\mathsf{d}_\mathcal{K}\big(\gamma_\varepsilon(s_0), \gamma_\varepsilon(s_1)\big)^2 \le |s_1 - s_0| M_{f_\varepsilon}^2 \int_{s_0}^{s_1} f_\varepsilon^2 \mathcal{A}(\gamma(s), \gamma'(s))\,\mathrm{d}s \le |s_1 - s_1|^2 M_{f_\varepsilon}^2,$$

which yields (3.31). □

Since in applications the metric $\mathcal{G}$ is often not given explicitly (see examples in Section 3.3), it is desirable to express the fundamental estimate (3.24) in terms of the Onsager operator $\mathcal{K} = \mathcal{G}^{-1}$.

Proposition 3.2.7 *Assume that*

$$\begin{aligned} &\forall\, u \in \mathcal{Z}\ \forall \eta \in \mathcal{G}(u)Y: \quad \langle \eta, \mathcal{M}(u)\eta\rangle \ge \lambda \langle \eta, \mathcal{K}(u)\eta\rangle, \\ &\textit{where } \langle \eta, \mathcal{M}(u)\eta\rangle \overset{\textit{def}}{=} \langle \eta, \mathrm{D}\mathcal{F}(u)\mathcal{K}(u)\eta\rangle - \tfrac{1}{2}\langle \eta, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\eta\rangle, \end{aligned} \tag{3.32}$$

then estimate (3.24) *holds.*

Proof: The proof is immediate since for a given $v \in Y$ we can use $\eta = \mathcal{G}(u)v$ in (3.32). After using the formula for the derivative of the inverse, namely $\mathrm{D}\mathcal{G}(u)[w] = -\mathcal{G}(u)\mathrm{D}\mathcal{K}(u)[w]\mathcal{G}(u)$ we find (3.24). □

Note that the conditions in Proposition 3.2.7 are closely related to the Bakry-Émery conditions [BaÉ85, Bak94] and provide a strengthened version of the classical entropy-dissipation estimate. In fact, defining the quantities $\mathcal{D}(u) = \langle \mathrm{D}\mathcal{E}(u), \mathcal{K}(u)\mathrm{D}\mathcal{E}(u)\rangle$ and $\mathcal{R}(u) = 2\langle \mathrm{D}\mathcal{E}(u), \mathcal{M}(u)\mathrm{D}\mathcal{E}(u)\rangle$ the solutions u of $\dot{u} = -\mathcal{K}(u)\mathrm{D}\mathcal{E}(u)$ satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(u(t)) = -\mathcal{D}(u(t)) \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{D}(u(t)) = -\mathcal{R}(u(t)).$$

By (3.4) there exists $\alpha \ge \lambda$ such that $\mathcal{R}(u) - 2\alpha\mathcal{D}(u) = \mathcal{P}(u) \ge 0$ for all u. Assuming $\alpha > 0$, in [AM*01] the decay estimates

$$\mathcal{D}(u(t)) \le \mathrm{e}^{-2\alpha t}\mathcal{D}(u(0)) \quad \text{and} \quad \mathcal{E}(u(t)) - \mathcal{E}(u(\infty)) + \int_t^\infty \mathcal{P}(u(s))\,\mathrm{d}s = \frac{1}{2\alpha}\mathcal{D}(u(t))$$

are used to derive convergence for $t \to \infty$. We discuss further useful properties of the geodesic λ-convexity in Section 3.2.3, also if $\lambda < 0$.

We now return to the metric evolution in the larger space $\mathcal{X}$. For this, we assume that $\mathsf{d}_\mathcal{K}$ on $\mathcal{Z}$ can be extended to a metric on $\mathcal{X}$ such that

$$(\mathcal{X}, \mathsf{d}_\mathcal{K}) \text{ is a complete metric space.} \tag{3.33}$$

Moreover, $\mathcal{E} : \mathcal{Z} \to \mathbb{R}$ is assumed to have a lower semicontinuous extension $\overline{\mathcal{E}} : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ (with respect to the metric topology). Finally, $\mathcal{Z}$ is assumed to be dense, viz.

$$\forall\, u \in \mathcal{X} \text{ with } \overline{\mathcal{E}}(u) < \infty \ \exists\, u_n \in \mathcal{Z} : \quad \mathsf{d}_{\mathcal{K}}(u_n, u) \to 0 \text{ and } \mathcal{E}(u_n) \to \overline{\mathcal{E}}(u). \tag{3.34}$$

Using the Lipschitz continuity (3.16), there is a unique continuous extension $\overline{\mathcal{S}} : [0,\infty[\times \mathcal{X} \to \mathcal{X}$. Then, [DaS08, Thm. 3.3] provides the following result.

Theorem 3.2.8 *If* (3.33), (3.34) *and the assumptions of Theorem 3.2.5 hold, then the semiflow $\overline{\mathcal{S}}$ associated with the gradient system $(\mathcal{X}, \overline{\mathcal{E}}, \mathsf{d}_{\mathcal{K}})$ satisfies EVI$_\lambda$* (3.23) *and the Lipschitz continuity* (3.16) *with $(\mathcal{E}, \mathcal{S})$ replaced by $(\overline{\mathcal{E}}, \overline{\mathcal{S}})$. Moreover, $\overline{\mathcal{E}}$ is geodesically λ-convex on $\mathcal{X}$, i.e. for every arc-length parameterized geodesic curve $\gamma \in \mathrm{C}^0([0,1]; \mathcal{X})$ we have*

$$\overline{\mathcal{E}}(\gamma(s)) \leq (1-s)\, \overline{\mathcal{E}}(\gamma(0)) + s\, \overline{\mathcal{E}}(\gamma(1)) - \frac{\lambda}{2} s(1-s)\, \mathsf{d}_{\mathcal{K}}(\gamma(0), \gamma(1))^2 \text{ for } s \in [0,1]. \tag{3.35}$$

3.3 Examples

This section surveys possible applications of the abstract methods developed in the previous section to scalar equations as well as reaction-diffusion systems. In particular, we show geodesic λ-convexity of gradient structures $(X, \mathcal{E}, \mathcal{K})$ in a smooth setting by establishing the estimate $\langle \xi, \mathcal{M}(u)\xi \rangle \geq \lambda \langle \xi, \mathcal{K}(u)\xi \rangle$. In particular, we generalize the known results for scalar drift-diffusion equations (with conserved mass) to systems with reaction terms (non-conserved masses). The discussion of the corresponding metric spaces $(X, \mathsf{d}_{\mathcal{K}})$ is postponed to future research.

3.3.1 Pure reaction systems and Markov chains

In [Mie11b] an entropy gradient structure was established for general reaction systems of mass-action type that satisfy the detailed balance condition. We consider a vector $\boldsymbol{u} \in\,]0,\infty[^I$ of densities and R polynomial reactions

$$\dot{\boldsymbol{u}} = -\sum_{r=1}^{R} k^r(\boldsymbol{u}) \Big(\frac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\alpha^r}} - \frac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}} \Big) (\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r), \qquad \text{where } \boldsymbol{u}^{\boldsymbol{\alpha}^r} = \Pi_{i=1}^I u_i^{\alpha_i^r}. \tag{3.36}$$

Here, $\boldsymbol{w} \in\,]0,\infty[^I$ is the reference density, which is obviously a steady state and satisfies the detailed balance condition. Moreover, $k^r(\boldsymbol{u}) \geq 0$ is the reaction coefficient (normalized with respect to $\boldsymbol{w}$), and the vectors $\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r \in\,]0,\infty[^I$ are the stoichiometric vectors for the forward and backward reactions. Usually the entries are assumed to be nonnegative integers, but this is not necessary here. As was shown in [Mie11b] the gradient system $(]0,\infty[^I, E, \mathbb{K})$ with

$$E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i \log(u_i/w_i) \quad \text{and} \quad \mathbb{K}(\boldsymbol{u}) = \sum_{r=1}^{R} k^r(\boldsymbol{u}) \Lambda\big(\tfrac{\boldsymbol{u}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}^{\boldsymbol{\alpha}^r}}, \tfrac{\boldsymbol{u}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}^{\boldsymbol{\beta}^r}} \big) (\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r) \otimes (\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r)$$

gives (3.36). We find $\langle \boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \boldsymbol{\xi}\cdot\boldsymbol{M}(\boldsymbol{u})\boldsymbol{\xi}$, where $\boldsymbol{M}(\boldsymbol{u}) \in \mathbb{R}^{I\times I}$ is defined via

$$\boldsymbol{M}(\boldsymbol{u}) = \tfrac{1}{2}\Big(\mathbb{K}(\boldsymbol{u})\mathrm{D}\mathcal{F}(\boldsymbol{u})^{\mathsf{T}} + \mathrm{D}\mathcal{F}(\boldsymbol{u})\mathbb{K}(\boldsymbol{u}) - \mathrm{D}\mathbb{K}(\boldsymbol{u})[\mathcal{F}(\boldsymbol{u})]\Big),$$

see also [Mie11a]. Note that the vector field $\mathcal{F}(\boldsymbol{u}) = \mathbb{K}(\boldsymbol{u})\mathrm{D}E(\boldsymbol{u})$ is nonlinear and that the matrices $\mathbb{K}(\boldsymbol{u})$ and $\mathbb{M}(\boldsymbol{u})$ have no homogeneity or concavity properties, in general.

We want to study a few simple cases and discuss the possibility of geodesic λ-convexity. For $R = 1$ we drop the reaction number r and write $\boldsymbol{\gamma} = \boldsymbol{\alpha} - \boldsymbol{\beta}$ and $\boldsymbol{\varrho} = (u_i/w_i)_i$. Then, we can write

$$\begin{aligned}
&\mathcal{F}(\boldsymbol{u}) = \phi(\boldsymbol{u})\boldsymbol{\gamma} && \text{with } \phi(\boldsymbol{u}) = k(\boldsymbol{u})(\boldsymbol{\varrho}^{\boldsymbol{\alpha}} - \boldsymbol{\varrho}^{\boldsymbol{\beta}}),\\
&\mathbb{K}(\boldsymbol{u}) = \kappa(\boldsymbol{u})\boldsymbol{\gamma}\otimes\boldsymbol{\gamma} && \text{with } \kappa(\boldsymbol{u}) = k(\boldsymbol{u})\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}}),\\
&\boldsymbol{M}(\boldsymbol{u}) = m(\boldsymbol{u})\boldsymbol{\gamma}\otimes\boldsymbol{\gamma} && \text{with } m(\boldsymbol{u}) = \kappa(\boldsymbol{u})\mathrm{D}\phi(\boldsymbol{u})\cdot\boldsymbol{\gamma} - \tfrac{1}{2}\phi(\boldsymbol{u})\mathrm{D}\kappa(\boldsymbol{u})\cdot\boldsymbol{\gamma}.
\end{aligned}$$

The general case seems too difficult to be analyzed, hence we reduce to the case $k(\boldsymbol{u}) \equiv 1$. Introducing the matrix $V = \mathrm{diag}(1/u_i)_i$ we have $\mathrm{D}_{\boldsymbol{u}}(\boldsymbol{u}^{\boldsymbol{\alpha}})[\gamma] = \boldsymbol{u}^{\boldsymbol{\alpha}}\boldsymbol{\alpha}\cdot V\boldsymbol{\gamma}$, and after some elementary calculations involving the properties of the function Λ (see [Mie11a]) we find

$$m(\boldsymbol{u}) = \tfrac{1}{2}\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})\big(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}\boldsymbol{\alpha} - \boldsymbol{\varrho}^{\boldsymbol{\beta}}\boldsymbol{\beta} + \Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\boldsymbol{\alpha}-\boldsymbol{\beta})\big)\cdot V(\boldsymbol{\alpha}-\boldsymbol{\beta}).$$

For geodesic λ-convexity we have to show $m(\boldsymbol{u}) \geq \lambda\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})$ which after dividing by $\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})$ leads to the formula

$$\lambda = \tfrac{1}{2}\inf\Big\{\textstyle\sum_{i=1}^{I}\frac{(\alpha_i-\beta_i)}{w_i\varrho_i}\big[\boldsymbol{\varrho}^{\boldsymbol{\alpha}}\alpha_i - \boldsymbol{\varrho}^{\boldsymbol{\beta}}\beta_i + \Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\alpha_i-\beta_i)\big] \;:\; \boldsymbol{\varrho}\in\,]0,\infty[^I\Big\}.$$

In the special case where $\alpha_i\beta_i = 0$ for all i we find the simpler form

$$\lambda = \tfrac{1}{2}\inf\Big\{\textstyle\sum_{i=1}^{I}\frac{1}{w_i\varrho_i}\big(\alpha_i^2\boldsymbol{\varrho}^{\boldsymbol{\alpha}}+\beta_i^2\boldsymbol{\varrho}^{\boldsymbol{\beta}}+\Lambda(\boldsymbol{\varrho}^{\boldsymbol{\alpha}}, \boldsymbol{\varrho}^{\boldsymbol{\beta}})(\alpha_i^2+\beta_i^2)\big) \;:\; \boldsymbol{\varrho}\in\,]0,\infty[^I\Big\} \geq 0.$$

This formula applies to example (2.6) where $\boldsymbol{\alpha} = (1,1,0,0)^{\mathsf{T}}$ and $\boldsymbol{\beta} = (0,0,1,1)^{\mathsf{T}}$. Because of $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \geq 2$ the infimum is $\lambda = 0$ (by choosing $\boldsymbol{\varrho} = \varepsilon(1,1,1,1)$ and $\varepsilon \to 0$).

Example 3.3.1 *The annihilation-creation reaction modeling recombination and generation of electron-hole pairs in semiconductors, cf. [Gli08, GlG09] and Section 2.1.5, reads*

$$\dot{\boldsymbol{u}} = -(u_1u_2 - 1)\,(1,1)^{\mathsf{T}}, \quad \textit{where} \quad \boldsymbol{\alpha} = (1,1)^{\mathsf{T}} \textit{ and } \boldsymbol{\beta} = (0,0)^{\mathsf{T}}. \tag{3.37}$$

The formula yields $\lambda = \frac{1}{2}\inf\Big\{\big(\frac{1}{u_1}+\frac{1}{u_2}\big)\big(u_1u_2+\Lambda(u_1u_2,1)\big) \;:\; u_1, u_2 > 0\Big\} = \cosh(1) > 0.$

Discrete Markov chains can be seen as special reaction systems where only exchange reactions $X_i \rightleftharpoons X_j$ occur. The reaction system takes the form

$$\dot{\boldsymbol{u}} = \boldsymbol{R}(\boldsymbol{u}) = Q\boldsymbol{u}, \quad \text{where } Q_{ij} \geq 0 \text{ for } i \neq j \quad \text{and } \textstyle\sum_{i=1}^{I} Q_{ij} = 0. \tag{3.38}$$

We assume that there is a unique steady state $\boldsymbol{w}$ with $w_i > 0$ for all i (also called irreducibility). A much stronger assumption is the condition of detailed balance, which reads $Q_{ij}w_j = Q_{ji}w_i$ for $i, j = 1, ..., I$. According to [Mie11a, Maa11], (3.38) is induced by the gradient system $(X_{\text{Mkv}}, E_{\text{Mkv}}, \mathbb{K}_{\text{Mkv}})$, with $X_{\text{Mkv}} = \left\{ \boldsymbol{u} \in [0,1]^I \,:\, \sum_{i=1}^I u_i = 1 \right\}$, $E_{\text{Mkv}}(\boldsymbol{u}) = \sum_{i=1}^I u_i \log(u_i/w_i)$, and

$$\mathbb{K}_{\text{Mkv}}(\boldsymbol{u}) = \sum_{1 \le i < j \le I} Q_{ij} w_j \, \Lambda\big(\tfrac{u_i}{w_i}, \tfrac{u_j}{w_j}\big) \, (\boldsymbol{e}_i - \boldsymbol{e}_j) \otimes (\boldsymbol{e}_i - \boldsymbol{e}_j) \ \in \mathbb{R}^{I \times I},$$

where $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^I$ are the unit vectors. Moreover, it is shown in [Mie11a] that for all Markov chains there is a $\lambda \in \mathbb{R}$ such that $(E_{\text{Mkv}}, \mathbb{K}_{\text{Mkv}})$ is geodesically λ-convex. For special classes, like tridiagonal Q, explicit estimates for λ are obtained.

3.3.2 Scalar diffusion equation

We consider a bounded, convex domain $\Omega \subset \mathbb{R}^d$, $d \ge 1$, with smooth boundary. In Ω we are given the scalar diffusion equation

$$\dot{u} = \operatorname{div}(a(u)\nabla u) \quad \text{in } \Omega, \qquad \nabla u \cdot \nu = 0 \quad \text{on } \partial\Omega. \tag{3.39}$$

This equation is the gradient flow of the energy $\mathcal{E}$ with respect to the Onsager operator $\mathcal{K}$ given via

$$\mathcal{E}_0(u) = \int_\Omega E(u)\,\mathrm{d}x \qquad \text{and} \qquad \mathcal{K}(u)\xi = -\operatorname{div}(\mu(u)\nabla\xi),$$

where E and μ are such that $\mu(u)E''(u) = a(u)$ holds. In particular, we assume that $E, \mu \in \mathrm{C}^2(]0, \infty[)$ and the sign conditions

$$\mu(u) \ge 0, \quad \mu''(u) \le 0, \quad E''(u) > 0 \quad \text{ for all } u > 0.$$

The choice, $\mu(u) = u$ leads to the well-known Wasserstein case. The subscript "0" in $\mathcal{E}_0$ reflects that there is no potential energy. This case will be considered in the subsequent example.

We impose that solutions $u : \Omega \to \mathbb{R}$ of (3.39) are sufficiently smooth for given smooth initial conditions such that the assumptions of the last section for the semiflow $\mathcal{S}_t : u(0) \mapsto u(t)$ are satisfied.

In the following we slightly deviate from the setting in Section 3.2.4 in that we consider $\mathcal{Y}$ and $\mathcal{Z}$ to be open and connected subsets of affine spaces $u_*{+}Y$ and $u_*{+}Z$ where the shift is given by $u_* = 1/|\Omega|$. This modification allows us to extend our theory to the space of probability measures. More precisely, let $X = \mathsf{Meas}(\Omega)$ the space of Radon measures ρ (using that Ω is bounded all moments $\int_\Omega |x|^p \,\mathrm{d}\rho(x)$ are finite) and $\mathcal{X} = \mathsf{Prob}(\Omega)$ denotes the subset of probability measures such that

$$\mathcal{X} = \mathsf{Prob}(\Omega) = \{\rho \in X \,:\, \rho(\Omega) = 1 \text{ and } \rho \ge 0\}.$$

The results of Section 3.2.4 can be easily adapted to this case.

The quadratic form associated with the operator $\mathcal{K}$ defines in a natural way the spaces

$$H^* = \mathrm{H}^1_{\mathrm{av}}(\Omega) = \Big\{ \xi \in \mathrm{H}^1(\Omega) \,:\, \textstyle\int_\Omega \xi \,\mathrm{d}x = 0 \Big\}, \quad H = \mathrm{H}^{-1}_{0,\mathrm{av}}(\Omega) = (\mathrm{H}^1_{\mathrm{av}}(\Omega))^*.$$

Moreover, we choose $s \geq 4$ such that $s > 2 + d/2$ and define the spaces

$$\begin{aligned}
&Y = \Big\{ v \in \mathrm{H}^{s-2}(\Omega) \,:\, \textstyle\int_\Omega v \,\mathrm{d}x = 0 \text{ and } \nabla v \cdot \nu = 0 \text{ on } \partial\Omega \Big\},\\
&\mathcal{Y} = \{ u \in u_* {+} Y \,:\, \inf u > 0 \},\\
&Z = \mathrm{H}^s(\Omega) \cap Y,\\
&\mathcal{Z} = \{ u \in u_* {+} Z \cap \mathcal{Y} \,:\, \nabla\big(\mathrm{div}(a(u)\nabla u)\big) \cdot \nu = 0 \text{ on } \partial\Omega \}.
\end{aligned}$$

The boundary condition in the definition of the set $\mathcal{Z}$ is necessary to ensure that the semiflow satisfies $\mathcal{S} \in \mathrm{C}^1([0,\infty[\times\mathcal{Z}; Y)$. In particular, for a solution $t \mapsto u(t) \in \mathcal{Z}$ holds $\dot{u}(t) \in Y$. Obviously we have $Z \subset Y \subset X$ and $\mathcal{Y} \subset u_* + Y$ and $\mathcal{Z} \subset u_* + Z$ with dense embeddings. Moreover, the assumptions on the Sobolev index s yield the embeddings

$$Y \subset \mathrm{H}^2(\Omega) \cap \mathrm{C}_\mathrm{b}(\Omega) \quad \text{and} \quad Z \subset \mathrm{C}^2_\mathrm{b}(\Omega), \tag{3.40}$$

where $\mathrm{C}_\mathrm{b}(\Omega)$ denotes the space of bounded continuous functions on Ω.

Our analysis is similar to that in [DaS08, Sect. 4] with the main difference that we have to take care of the boundary conditions when doing integrations by part. There are two crucial observations for the case with boundaries: Firstly, the curvature of the boundary of convex bodies provides a sign for the normal derivative $\nabla(|\nabla\xi|^2) \cdot \nu \leq 0$, whenever $\nabla\xi \cdot \nu = 0$ holds, see Proposition 3.3.2. Secondly, the test functions $\xi \in \mathcal{G}(u)Y$ will satisfy *two boundary conditions*, namely

$$-\mathrm{div}(\mu(u)\nabla\xi) = v \in Y \quad \Longrightarrow \quad \big(\, \nabla\xi \cdot \nu = 0 \;\text{ and }\; \nabla\big(\mathrm{div}(\mu(u)\nabla\xi)\big) \cdot \nu = 0 \,\big).$$

In order to show the geodesic λ-convexity of $\mathcal{E}_0$ with respect to $\mathcal{K}$ we prove that the assumptions of Proposition 3.2.7 hold. We have to compute the quadratic form $\langle \xi, \mathcal{M}_0(u)\xi \rangle = \langle \xi, \mathrm{D}\mathcal{F}_0(u)\mathcal{K}(u)\xi \rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}_0(u)]\xi \rangle$, with

$$\mathcal{F}_0(u) = -\mathrm{div}(a(u)\nabla u) \quad \text{and} \quad \mathrm{D}\mathcal{F}_0(u)[v] = -\mathrm{div}\big(a'(u)v\nabla u + a(u)\nabla v\big).$$

For $\xi \in \mathcal{G}(u)Y$ we use the abbreviation $v = \mathcal{K}(u)\xi \in Y$ and obtain by integration by parts

$$\begin{aligned}
\langle \xi, \mathcal{M}_0(u)\xi \rangle &= -\int_\Omega \xi\big(\mathrm{div}(a'(u)v\nabla u{+}a(u)\nabla v)\big)\,\mathrm{d}x - \frac{1}{2}\int_\Omega \mu'(u)\big(-\mathrm{div}(a(u)\nabla u)\big)|\nabla\xi|^2\,\mathrm{d}x\\
&= \int_\Omega \nabla\xi \cdot \big(a'(u)v\nabla u + a(u)\nabla v\big)\,\mathrm{d}x - \int_\Omega a(u)\nabla u \cdot \nabla\big(\mu'(u)\tfrac{1}{2}|\nabla\xi|^2\big)\,\mathrm{d}x,
\end{aligned}$$

where in both cases the boundary terms vanish, namely using $(a'(u)v\nabla u{+}a(u)\nabla v) \cdot \nu = 0$ and $a(u)\nabla u \cdot \nu = 0$ from $v \in Y$ and $u \in \mathcal{Z}$. Moreover, all integrals above are welldefined due to (3.40).

Applying integration by parts one more time yields

$$\begin{aligned}\langle \xi, \mathcal{M}_0(u)\xi\rangle &= \int_\Omega a(u)\Delta\xi \operatorname{div}(\mu(u)\nabla\xi)\,\mathrm{d}x - \int_\Omega a(u)\nabla u\cdot\nabla\big(\mu'(u)\tfrac12|\nabla\xi|^2\big)\,\mathrm{d}x\\ &= \int_\Omega \nabla H(u)\cdot(\Delta\xi\nabla\xi - \nabla(\tfrac12|\nabla\xi|^2)) + a(u)\mu(u)(\Delta\xi)^2 - \frac{a(u)\mu''(u)}{2}|\nabla u|^2|\nabla\xi|^2\,\mathrm{d}x,\end{aligned}$$

where we have set $H(u) = \int_0^u a(y)\mu'(y)\,\mathrm{d}y$ and used that $\nabla\xi\cdot\nu = 0$. Finally, integrating by parts one last time leads to

$$\begin{aligned}\langle \xi, \mathcal{M}_0(u)\xi\rangle = \int_\Omega & H(u)|\mathrm{D}^2\xi|^2 + \big(a(u)\mu(u) - H(u)\big)(\Delta\xi)^2 - \frac{a(u)\mu''(u)}{2}|\nabla u|^2|\nabla\xi|^2\,\mathrm{d}x\\ &- \int_{\partial\Omega} H(u)\nabla(\tfrac12|\nabla\xi|^2)\cdot\nu\,\mathrm{da}.\end{aligned} \tag{3.41}$$

Here, we used Bochner's formula $\operatorname{div}((\Delta\xi)\nabla\xi) - \Delta(\frac12|\nabla\xi|^2) = (\Delta\xi)^2 - |\mathrm{D}^2\xi|^2$. We observe that the boundary integral is nonpositive using the assumption $H(u)\geq 0$ and Proposition 3.3.2 below.

Thus, we have shown that $\langle\xi,\mathcal{M}_0(u)\xi\rangle \geq 0$ holds if we assume that $u\mapsto\mu(u)$ is concave and $a\mu \geq \frac{d-1}{d}H \geq 0$. Here, the latter condition is due to the elementary estimate

$$\forall \mathbb{E}\in\mathbb{R}^{d\times d}: \qquad \alpha|\mathbb{E}|^2 - \beta(\operatorname{tr}\mathbb{E})^2 \geq 0 \quad\Leftrightarrow\quad \alpha \geq \max\{0, d\beta\}.$$

Now, Proposition 3.2.7 states that $(\mathcal{E}_0,\mathcal{K})$ is geodesically 0-convex. Since the present result will be a special case of the result in the next subsection, we refer to Theorem 3.3.3 for the precise statement.

Thus, we have generalized [DaS08, Thm. 4.2] from manifolds without boundary to the case of convex domains in $\mathbb{R}^d$ with smooth boundaries. The condition of convexity is quite natural in the context of optimal transport, since only convex domains are still complete metric length-spaces with respect to the Euclidean distance.

We used the following proposition on the signs of $\nabla\big(|\nabla\xi|^2\big)\cdot\nu$ on the boundary. We refer to [Gri85, Ch. 3] and [GST09, Lem. 5.2] for previous proofs, but still give an independent proof of a more general result needed in Section 3.3.7. It involves the second fundamental form $\mathbb{I}$ of the boundary, i.e. for two tangent vectors $\tau_1,\tau_2\in \mathrm{T}_x\partial\Omega$ we have $\mathbb{I}(\tau_1,\tau_2) = \tau_1\cdot\mathrm{D}\nu(x)\tau_2 = \mathbb{I}(\tau_2,\tau_1)$, where ν is the outer normal vector.

Proposition 3.3.2 *Assume that $\Omega\subset\mathbb{R}^d$ is a domain with C^2 boundary. Then, for functions $\xi_1,\xi_2\in\mathrm{H}^3(\Omega)$ with $\nabla\xi_1\cdot\nu = \nabla\xi_2\cdot\nu = 0$ on $\partial\Omega$ we have the identity*

$$\nabla\big(\nabla\xi_1\cdot\nabla\xi_2\big)\cdot\nu = -2\mathbb{I}(\nabla_{\|}\xi_1,\nabla_{\|}\xi_2), \tag{3.42}$$

where $\nabla_{\|}\xi$ denotes the tangential part of the gradient $\nabla_{\|}\xi = \nabla\xi - (\nabla\xi\cdot\nu)\nu$. In particular, if Ω is convex and $\xi_2 = \xi_1$, then $\nabla\big(|\nabla\xi_1|^2\big)\cdot\nu \leq 0$ on $\partial\Omega$.

Proof: Without loss of generality we assume that ξ_j is smooth. We denote by $\overline{\nu}\in\mathrm{C}^1(\overline{\Omega})$

a smooth extension of the outer unit normal ν into Ω. For $x \in \Omega$ we compute

$$\begin{aligned} \nabla\Big(\nabla\xi_1 \cdot \nabla\xi_2\Big) \cdot \overline{\nu}(x) &= \nabla\xi_2 \cdot \mathrm{D}^2\xi_1\,\overline{\nu} + \nabla\xi_1 \cdot \mathrm{D}^2\xi_2\,\overline{\nu} \\ &= \nabla\xi_2 \cdot \Big(\nabla(\nabla\xi_1 \cdot \overline{\nu}) - \mathrm{D}\overline{\nu}\nabla\xi_1\Big) + \nabla\xi_1 \cdot \Big(\nabla(\nabla\xi_2 \cdot \overline{\nu}) - \mathrm{D}\overline{\nu}\nabla\xi_2\Big). \end{aligned} \tag{3.43}$$

On the boundary the product $\nabla\xi_j \cdot \overline{\nu}$ vanishes identically, such that $\nabla_{\|}(\nabla\xi_j \cdot \overline{\nu}) = 0$ on $\partial\Omega$. Hence, there are scalar functions $\gamma_j : \partial\Omega \to \mathbb{R}$ such that $\nabla(\nabla\xi_j \cdot \overline{\nu}) = \gamma\nu$ on $\partial\Omega$. Inserting this into (3.43) and using $\nabla\xi_j \cdot \nu = 0$ we have established (3.42).

For a convex body, the second fundamental form is positive semidefinite. Hence, formula (3.42) gives the desired result for $\xi_1 = \xi_2$. □

We end this subsection by mentioning that the theory can also be applied to smooth inhomogeneous systems, e.g. where the mobility depends on the spatial variable $x \in \Omega$:

$$\begin{aligned} &\dot{u} = -\mathcal{F}(u) = \mathrm{div}(\mathbb{M}(x)\nabla u), \qquad \text{with} \\ &\mathcal{K}(u)\xi = -\mathrm{div}(u\,\mathbb{M}(x)\nabla\xi), \quad \mathcal{E}(u) = \int_\Omega u\log u\,\mathrm{d}x, \end{aligned}$$

where $\mathbb{M} \in \mathrm{W}^{2,\infty}(\overline{\Omega}; \mathbb{R}^{d\times d}_{\mathrm{spd}})$, and there exists $\alpha_0 > 0$ with $\boldsymbol{a}\cdot\mathbb{M}(x)\boldsymbol{a} \geq \alpha_0|\boldsymbol{a}|^2$.

The appropriate boundary conditions are now $(\mathbb{M}(x)\nabla u(x)) \cdot \nu(x) = 0 = (\mathbb{M}(x)\nabla\xi(x)) \cdot \nu(x)$ for $x \in \partial\Omega$. Doing the appropriate integrations by part we obtain the formula

$$\begin{aligned} &\langle\xi, \mathcal{M}(u)\xi\rangle = \langle \mathrm{D}\mathcal{F}(u)\xi, \mathcal{K}(u)\xi\rangle - \tfrac{1}{2}\langle\xi, \mathrm{D}\mathcal{K}(u)[\mathcal{F}(u)]\xi\rangle \\ &= \int_\Omega \mathrm{div}(\mathbb{M}\nabla\xi)\mathrm{div}(u\mathbb{M}\nabla\xi) + \tfrac{1}{2}\mathrm{div}(\mathbb{M}\nabla u)\nabla\xi \cdot \mathbb{M}\nabla\xi\,\mathrm{d}x \\ &= \int_\Omega u\Big(\nabla\xi\cdot\mathbb{B}\nabla\xi + \nabla\xi\cdot\boldsymbol{B}{:}\mathrm{D}^2\xi + |\mathbb{M}\mathrm{D}^2\xi|^2\Big)\,\mathrm{d}x - \int_{\partial\Omega} u\mathbb{M}\nabla\big(\tfrac{1}{2}\nabla\xi\cdot\mathbb{M}\nabla\xi\big) \cdot \nu\,\mathrm{d}a, \end{aligned}$$

where all terms involving third derivatives of ξ cancel, and the tensors $\mathbb{B}$ and $\boldsymbol{B}$ are given via $\mathbb{M}$, $\mathrm{D}\mathbb{M}$, and $\mathrm{D}^2\mathbb{M}$. Proposition 3.3.2 can be generalized for spatially dependent mobilities leading to three additional terms due to the spatial derivatives of $\mathbb{M}$:

$$\begin{aligned} \mathbb{M}\nabla\big(\nabla\xi_1\cdot\mathbb{M}\nabla\xi_2\big) \cdot \nu = -\mathbb{I}(\mathbb{M}\nabla\xi_1, \mathbb{M}\nabla\xi_2) + \nabla\xi_1 \cdot \mathrm{D}\mathbb{M}[\mathbb{M}\nu]\nabla\xi_2 \\ - \nabla\xi_2 \cdot \mathrm{D}\mathbb{M}[\mathbb{M}\nabla\xi_1]\nu - \nabla\xi_1 \cdot \mathrm{D}\mathbb{M}[\mathbb{M}\nabla\xi_2]\nu. \end{aligned}$$

If the sum of these terms is negative, using $\alpha_0 > 0$ and $\mathbb{M} \in \mathrm{W}^{2,\infty}(\Omega)$ (giving $\boldsymbol{B}, \mathbb{B} \in \mathrm{L}^\infty(\Omega)$) and pointwise minimization over $\mathrm{D}^2\xi(x) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ provides a $\lambda_{\mathbb{M}} \in \mathbb{R}$ such that $\langle\xi, \mathcal{M}(u)\xi\rangle \geq \lambda_{\mathbb{M}}\langle\xi, \mathcal{K}(u)\xi\rangle = \lambda_{\mathbb{M}}\int_\Omega u\,\nabla\xi\cdot\mathbb{M}\nabla\xi\,\mathrm{d}x$.

For an isotropic mobility matrix $\mathbb{M}(x) = \mu(x)I$, that satisfies the boundary relation

$\mu\mathbb{I}(\tau,\tau) \geq \nabla\mu\cdot\nu|\tau|^2$ for all $x \in \partial\Omega$ and $\tau \in \mathrm{T}_x\partial\Omega$, we obtain the simplified estimate

$$\begin{aligned}\langle \xi, \mathcal{M}(u)\xi\rangle &= \int_\Omega u\Big(\nabla\xi\cdot((\tfrac{1}{2}\mu\Delta\mu+\tfrac{1}{2}|\nabla\mu|^2)I - \mu\mathrm{D}^2\mu)\nabla\xi \\ &\qquad + 2\mu\nabla\xi\cdot\mathrm{D}^2\xi\nabla\mu - \mu\Delta\xi\nabla\mu\cdot\nabla\xi + \mu^2|\mathrm{D}^2\xi|^2\Big)\,\mathrm{d}x \\ &\quad + \int_{\partial\Omega} u(\mu^2\mathbb{I}(\nabla\xi,\nabla\xi) - \mu(\nabla\mu\cdot\nu)|\nabla\xi|^2)\,\mathrm{d}x \\ &\geq \int_\Omega u\Big(\nabla\xi\cdot(\tfrac{\mu}{2}\Delta\mu I - \mu\mathrm{D}^2\mu)\nabla\xi - \tfrac{d-2}{4}(\nabla\xi\cdot\nabla\mu)^2\Big)\,\mathrm{d}x \ \geq \lambda_{\mathbb{M}}\langle\xi,\mathcal{K}(u)\xi\rangle\end{aligned} \tag{3.44}$$

with $\lambda_{\mathbb{M}} = \inf\Big\{\frac{1}{2}\Delta\mu(x) - \sigma_{\max}(\mathrm{D}^2\mu(x) + \frac{d-2}{4\mu(x)}\nabla\mu(x)\otimes\nabla\mu(x)) \,:\, x\in\Omega\Big\}$, where again minimization with respect to $\mathrm{D}^2\xi$ is used in the first estimate. Here, $\sigma_{\max}(H) \in \mathbb{R}$ denotes the largest eigenvalue of a symmetric matrix $H \in \mathbb{R}^{d\times d}$. In space dimensions $d = 1$ and 2 we obtain

$$d=1: \quad \lambda_{\mathbb{M}} = \inf\Big\{-\mu''(x)/2 + (\mu'(x))^2/(4\mu(x)) \,:\, x\in\Omega\Big\}, \tag{3.45a}$$

$$d=2: \quad \lambda_{\mathbb{M}} = \inf\Big\{\tfrac{1}{2}(\sigma_{\min}(\mathrm{D}^2\mu(x)) - \sigma_{\max}(\mathrm{D}^2(\mu(x))) \,:\, x\in\Omega\Big\}. \tag{3.45b}$$

Our result can be compared to the estimates obtained in [Lis09, Thm. 1.5] with complete different methods. The results there are formulated using the Wasserstein distance $\mathsf{d}_{\mathrm{Wass}} = W_I$, while our results are formulated in terms of $\mathsf{d}_{\mathcal{K}}$ which is called W_G there, where $G(x) = \mathbb{M}^{-1}(x)$ (see [Lis09, Eqn. (1.67)]). Thus, our rate $\lambda_{\mathbb{M}}$ may differ from the contractivity rate α, which takes the form

$$\alpha = \inf\Big\{-\sigma_{\max}(\mathrm{D}^2\mu(x) + d\,\nabla\sqrt{\mu(x)}\otimes\nabla\sqrt{\mu(x)}) \,:\, x\in\Omega\Big\}$$

in our smooth setting.

3.3.3 A scalar drift-diffusion equation with concave mobility

We now generalize the diffusion equation of the previous section by adding a drift term induced by a given potential V. Moreover, we allow the density u to be restricted to a bounded interval, i.e. we assume that there is a bound $U \in\,]0,\infty]$ such that

$$u(t,x) \in\,]0, U[\quad \text{holds for almost every } t \text{ and } x.$$

Such restrictions occur in systems with exclusion principles. We refer to [GaG05, BD*10] and Section 3.3.7. Our work relates to [CL*10] and [LMS12, Prop. 4.6], where the entropy and the potential energy are studied concerning their geodesic λ-convexity. We make the result of the latter work more precise. We have the total energy and the Onsager operator

$$\mathcal{E}_V(u) = \int_\Omega E(u(x)) + u(x)V(x)\,\mathrm{d}x \quad \text{and} \quad \mathcal{K}(u)\xi = -\mathrm{div}(\mu(u)\nabla\xi).$$

The drift-diffusion equation takes the form

$$\dot{u} = \operatorname{div}(a(u)\nabla u + \mu(u)\nabla V) \quad \text{in } \Omega, \quad (a(u)\nabla u + \mu(u)\nabla V)\cdot\nu = 0 \quad \text{on } \partial\Omega, \tag{3.46}$$

where $a(u) = \mu(u)E''(u)$. We again impose the sign conditions

$$\mu(u) > 0, \quad \mu''(u) \le 0, \quad E''(u) > 0 \quad \text{for all } u \in\,]0, U[. \tag{3.47}$$

In the case $U < \infty$ we explicitly allow for the case $\mu'(u) < 0$ which occurs in the commonly used mobility $\mu(u) = u - u^2$ on $]0,1[$. We will see that the non-monotonicity of μ gives rise to new conditions. We emphasize that the following result does not need the condition $\nabla V \cdot \nu = 0$ on $\partial\Omega$ employed in [LMS12, Prop. 4.6].

Theorem 3.3.3 *Assume that Ω is a convex bounded domain in $\mathbb{R}^d$ with smooth boundary. In addition to* (3.47) *define $H(u) = \int_0^u \mu(y)\mu'(y)E''(y)\,\mathrm{d}y$ and assume*

$$H(u) \ge 0, \quad \mu(u)^2 E''(u) \ge \tfrac{d-1}{d} H(u) \quad \textit{for all } u \in\,]0, U[. \tag{3.48}$$

If the potential $V : \Omega \to \mathbb{R}$ satisfies $V \in \mathrm{W}^{2,\infty}(\Omega)$, then $(\mathcal{E}_V, \mathcal{K})$ are geodesically λ-convex for $\lambda = \lambda_2^V - \lambda_1^V$, where

$$\begin{aligned} \lambda_1^V &= \frac{9}{8}\|\nabla V\|_{\mathrm{L}^\infty}^2 \sup\left\{-\mu''(u)/E''(u) \,:\, u \in\,]0,U[\right\} \ge 0, \\ \lambda_2^V &= \inf\left\{\mu'(u)\,\boldsymbol{a}\cdot \mathrm{D}^2V(x)\boldsymbol{a} \,:\, u \in\,]0,U[, x \in \Omega, \boldsymbol{a} \in \mathbb{R}^d \textit{ with } |\boldsymbol{a}| = 1\right\}. \end{aligned}$$

Before giving the proof of this result note that the case of a linear mobility (i.e. $\mu(u) = u$) for the Wasserstein distance gives the standard result as $\lambda_1^V = 0$. Moreover, λ_2^V simply characterizes the λ-convexity of V on the Euclidean space Ω. Note that in the case $\mu'(u) < 0$ we need λ-concavity of V.

Proof: We proceed exactly as in the previous subsection. We only have the new terms associated with V. Since $\mathcal{K}$ is independent of V and the vector field $\mathcal{F}$ depends linearly on V, the new terms are also linear in V. Together with $\mathcal{M}_0$ from (3.41) we have

$$\begin{aligned} \langle \xi, \mathcal{M}_V(u)\xi\rangle &= \langle \xi, \mathcal{M}_0(u)\xi\rangle \\ &\quad + \int_\Omega \mu\mu'\nabla\xi\cdot\mathrm{D}^2V\nabla\xi + \frac{\mu\mu''}{2}\big(2\nabla u\cdot\nabla\xi\;\nabla V\cdot\nabla\xi - |\nabla\xi|^2\nabla V\cdot\nabla u\big)\,\mathrm{d}x. \end{aligned} \tag{3.49}$$

To reach this result, we emphasize that the integrations by parts have to be done of the full vector field $\mathcal{F}_V$ such that $w = \mathcal{K}(u)\xi$ in $\int_\Omega \xi \mathrm{D}\mathcal{F}_V(u)[w]\,\mathrm{d}x$ satisfies the additional boundary condition $[w(a'(u)\nabla u{+}\mu'(u)\nabla V) + a(u)\nabla w]\cdot\nu = 0$ obtained by differentiating the boundary condition in (3.46).

While the first term in (3.49) can be immediately estimated from below by $\lambda_2^V\mu|\nabla\xi|^2$, the other terms do not have a sign. That is why in [CL*10] it was expected that the potential energy $\int_\Omega uV\,\mathrm{d}x$ is not geodesically convex. However, to estimate the geodesic convexity of $\mathcal{E}$ we can use the nonnegative term $-\mu''\frac{a}{2}|\nabla u|^2|\nabla\xi|^2$ occurring in $\mathcal{M}_0$ and not needed otherwise to show positivity of $\mathcal{M}_0$. Abbreviating $\boldsymbol{U} = \nabla u$ and $\boldsymbol{X} = \nabla\xi$, we have to estimate the following terms from below:

$$
\begin{aligned}
&-\mu''\tfrac{a}{2}|\boldsymbol{U}|^2|\boldsymbol{X}|^2 + \tfrac{\mu\mu''}{2}\big(2\boldsymbol{U}\cdot\boldsymbol{X}\ \nabla V\cdot\boldsymbol{X} - |\boldsymbol{X}|^2\nabla V\cdot\boldsymbol{U}\big) \geq (-\mu'')|\boldsymbol{X}|^2\big(\tfrac{a}{2}|\boldsymbol{U}|^2 - \tfrac{3}{2}\mu|\boldsymbol{U}||\nabla V|\big)\\
&\geq \tfrac{9\mu''\mu^2}{8a}|\boldsymbol{X}|^2|\nabla V|^2 = \tfrac{9\mu''\mu}{8E''}|\nabla V|^2|\boldsymbol{X}|^2 \geq -\lambda_1^V\mu|\boldsymbol{X}|^2.
\end{aligned}
$$

Thus, the result is established. $\square$

We conclude by making the conditions more explicit in the case of $\mu(u) = u - u^2$ on $]0,1[$ and $E''(u) = 1/\mu(u)$, i.e. $E(u) = u\log u + (1-u)\log(1-u)$. We obtain $\lambda_1^V = 9\|\nabla V\|_\infty^2/16$ and $\lambda_2^V = \|r_{\text{spec}}(\mathrm{D}^2V(\cdot))\|_\infty$, where r_{spec} denotes the spectral radius.

3.3.4 A scalar nonlinear reaction-diffusion equation

In a convex, bounded, and smooth domain Ω we consider the reaction-diffusion equation

$$\dot{u} = \Delta u - f(u) \quad \text{in } \Omega, \qquad \nabla u\cdot\nu = 0 \quad \text{on } \partial\Omega.$$

We assume that it is the gradient flow of the free energy $\mathcal{E}$ and the Onsager operator $\mathcal{K}$ defined via

$$\mathcal{E}(u) = \int_\Omega u(\log u - 1)\,\mathrm{d}x \qquad \text{and} \qquad \mathcal{K}(u)\xi = -\operatorname{div}(u\nabla\xi) + \kappa(u)\xi. \tag{3.50}$$

Hence, we assume the relation $f(u) = \kappa(u)\log u$. The reaction coefficient κ satisfies

$$
\begin{aligned}
&\kappa \in \mathrm{C}^0([0,\infty[) \cap \mathrm{C}^2(]0,\infty[),\\
&\kappa(0) = 0, \quad \kappa(u), \kappa'(u) > 0 \text{ and } \kappa''(u) \leq 0 \quad \text{for all } u > 0.
\end{aligned} \tag{3.51}
$$

The concavity of κ implies that of $u \mapsto \langle\xi, \mathcal{K}(u)\eta\rangle$, which is the prerequisite of the convexity of $\mathsf{d}_{\mathcal{K}}^2$, see Remark 3.2.1.

Similar to the previous examples we introduce the spaces

$$
\begin{aligned}
&H^* = \mathrm{H}^1(\Omega), && H = \mathrm{H}_0^{-1}(\Omega),\\
&Y = \Big\{u \in \mathrm{H}^{s-2}(\Omega) \,:\, \nabla u\cdot\nu = 0 \text{ on } \partial\Omega\Big\}, && \mathcal{Y} = \{u \in Y \,:\, \inf u > 0\},\\
&Z = \mathrm{H}^s(\Omega) \cap Y, && \mathcal{Z} = \{u \in Z \cap \mathcal{Y} \,:\, \nabla(\Delta u - f(u))\cdot\nu = 0\}
\end{aligned}
$$

and calculate the quadratic form $\langle\xi, \mathcal{M}(u)\xi\rangle$. With $\mathrm{D}\mathcal{F}(u)[v] = -\Delta v - f'(u)v$ and $v = \mathcal{K}(u)\xi \in Y$ we obtain

$$\langle\xi, \mathcal{M}(u)\xi\rangle = \int_\Omega \xi\big(-\Delta v + f'(u)v\big)\,\mathrm{d}x - \frac{1}{2}\int_\Omega \big(-\Delta u + f(u)\big)\big(|\nabla\xi|^2 + \kappa'(u)\xi^2\big)\,\mathrm{d}x = I_1 + I_2.$$

Integrating twice the first term in I_1 (using the boundary conditions $\nabla\xi\cdot\nu = \nabla v\cdot\nu = 0$

on $\partial\Omega$) and inserting the definition of $v = \mathcal{K}(u)\xi$ we find

$$\begin{aligned} I_1 &= \int_\Omega \big(-\Delta\xi + f'(u)\xi\big)\,\big(-\mathrm{div}(u\nabla\xi) + \kappa(u)\xi\big)\,\mathrm{d}x \\ &= \int_\Omega -u\nabla\Delta\xi\cdot\nabla\xi + u\nabla\big(f'(u)\xi\big)\cdot\nabla\xi + \nabla\xi\cdot\nabla\big(\kappa(u)\xi\big) + f'(u)\kappa(u)\xi^2\,\mathrm{d}x \\ &= \int_\Omega -u\nabla\Delta\xi\cdot\nabla\xi + (uf'{+}\kappa)|\nabla\xi|^2 + (uf''{+}\kappa')\xi\nabla\xi{\cdot}\nabla u + f'\kappa\xi^2\,\mathrm{d}x. \end{aligned}$$

Similarly, we integrate by parts the first term in I_2 (using $\nabla u\cdot\nu = 0$) and obtain

$$\begin{aligned} 2I_2 &= \int_\Omega -\nabla u\cdot\nabla\big(|\nabla\xi|^2{+}\kappa'\xi^2\big) - f|\nabla\xi|^2 - 2f\kappa'\xi^2\,\mathrm{d}x \\ &= \int_\Omega u\Delta\big(|\nabla\xi|^2\big){-}f|\nabla\xi|^2{-}2\kappa'\xi\nabla\xi{\cdot}\nabla u{-}\big(\kappa''|\nabla u|^2{+}f\kappa'\big)\xi^2\,\mathrm{d}x \\ &\quad - \int_{\partial\Omega} u\nabla\big(|\nabla\xi|^2\big){\cdot}\nu\,\mathrm{d}a. \end{aligned}$$

Again using Proposition 3.3.2 we can estimate the boundary integral and obtain

$$\langle\xi,\mathcal{M}(u)\xi\rangle \geq \int_\Omega u|\mathrm{D}^2\xi|^2 + M_1(u)|\nabla\xi|^2 + m_2(u)\xi\,\nabla\xi{\cdot}\nabla u + \big(M_3(u){-}\kappa''|\nabla u|^2/2\big)\xi^2\,\mathrm{d}x$$

where $m_2(u) = uf''(u)$ and the auxiliary functions M_1, M_3 are defined as

$$M_1(u) = uf'(u) + \kappa(u) - f(u)/2, \quad M_3(u) = f'(u)\kappa(u) - f(u)\kappa'(u)/2.$$

Using assumption (3.51) the last term, which involves $|\nabla u|^2\xi^2$ is nonnegative and can be dropped. We define $M_2(u) = f(0) + uf'(u) - f(u)$ such that $M_2'(u) = m_2(u)$ and $M_2(0) = 0$. The term involving m_2 can be integrated by parts (using $\nabla\xi\cdot\nu = 0$) via

$$\int_\Omega M_2'(u)\nabla u\cdot\xi\nabla\xi\,\mathrm{d}x = -\int_\Omega M_2(u)\Delta\big(\tfrac12\xi^2\big)\,\mathrm{d}x = -\int_\Omega M_2(u)\big(|\nabla\xi|^2{+}\xi\Delta\xi\big)\,\mathrm{d}x.$$

The pointwise estimate $-M_2(u)\xi\Delta\xi \geq -\frac{u}{d}(\Delta\xi)^2 - \frac{dM_2(u)^2}{4u}\xi^2$ yields the lower estimate

$$\langle\xi,\mathcal{M}(u)\xi\rangle \geq \int_\Omega u\big(|\mathrm{D}^2\xi|^2{-}\tfrac{1}{d}(\Delta\xi)^2\big) + \big(M_1(u){-}M_2(u)\big)|\nabla\xi|^2 + \Big(M_3(u){-}\tfrac{dM_2(u)^2}{4u}\Big)\xi^2\,\mathrm{d}x.$$

Thus, we have established the following result.

Theorem 3.3.4 *Let Ω, κ, and M_j be given as above. Define the values*

$$\lambda_1^* = \inf\Big\{\tfrac{M_1(u)-M_2(u)}{u}\ :\ u>0\Big\} \quad \textit{and} \quad \lambda_2^* = \inf\Big\{\tfrac{4uM_3(u)-dM_2(u)^2}{4u\kappa(u)}\ :\ u>0\Big\},$$

and set $\lambda^* = \min\{\lambda_1^*,\lambda_2^*\}$. If $\lambda^* > -\infty$, then $(\mathcal{E},\mathcal{K})$ defined in (3.50) *is geodesically λ^*-convex.*

Proof: To conclude the proof we have to establish

$$\langle \xi, \mathcal{M}(u)\xi\rangle \geq \lambda^* \langle \xi, \mathcal{K}(u)\xi\rangle = \lambda^* \int_\Omega u|\nabla\xi|^2 + \kappa(u)\xi^2 \,\mathrm{d}x$$

for all $u \in \mathcal{Z}$ and $\xi \in \mathcal{G}(u)Y$. Since the first term in the above lower estimate for $\mathcal{M}$ is nonnegative, it suffices to show $M_1(u)-M_2(u) \geq \lambda^* u$ and $M_3(u)-dM_2(u)^2/(4u) \geq \lambda^*\kappa(u)$ for all $u \geq 0$. Since these estimates are exactly the definitions of λ_j^*, the desired result is established. $\square$

The following result provides sufficient conditions on the function κ, satisfying (3.51), that lead to a geodesically λ-convex gradient system. It is posed in terms of the ansatz $\kappa(u) = k(u)\Lambda(1,u)$ and shows that k can be chosen to be constant near $u = 0$ given the linear reaction term $f(u) = k(0)(u-1)$ there. For large u one may choose $k(u) = c(\log u)^p$ for $c > 0$ and $p \in [0,1]$ leading to the nonlinear reaction term $f(u) = c(u-1)(\log u)^{p-1}$.

Proposition 3.3.5 *Consider a function κ satisfying* (3.51) *and let $k(u) = \kappa(u)/\Lambda(1,u)$ be strictly positive. If there exist $0 < u_0 < 1 < u_1 < \infty$ and positive constants k_j, $j = 0,\dots,3$ such that k satisfies the conditions*

$$k \in \mathrm{C}^0([0,\infty[) \text{ with } k(0) = k_0; \tag{3.52a}$$

$$\liminf_{u\to\infty} k(u) \geq k_1; \tag{3.52b}$$

$$k \in \mathrm{C}^1([u_1,\infty[) \text{ and } k \in \mathrm{C}^{1,\alpha}([0,u_0]) \text{ for some } \alpha \in\,]1/2,1]\,; \tag{3.52c}$$

$$k(u)+uk'(u) \geq k_2 \text{ and } |k(u)+u^2k'(u)|^2 \leq k_3u^2k(u)/\log u \text{ for } u \geq u_1, \tag{3.52d}$$

then in Theorem 3.3.4 we have $\lambda^* > -\infty$. The case $k \equiv k_0$ gives $\lambda^* = \frac{k_0}{2}$.

Proof: We denote by $\eta_j(u)$ the functions in the infima defining λ_j^* in Theorem 3.3.4. Since both functions are continuous on $]0,\infty[$ it suffices to estimate η_j near $u = 0$ and $u = \infty$.

ad η_1: By (3.52a) we have $M_1(0)-M_2(0) = -f(0)/2 = k_0/2 > 0$ and conclude $\eta_1(u) \geq 0$ for sufficiently small u. For $u \geq 2$ we have

$$M_1(u) - M_2(u) = \kappa(u) + f(u)/2 - f(0) \geq \tfrac{\kappa(u)}{2}\log u = \tfrac{u-1}{2}k(u) \geq uk(u)/4.$$

Using (3.52b) we obtain $\eta_1(u) \geq k_1/4$ for all sufficiently large u.

ad η_2: For $u \leq 1$ we have $f(u) \leq 0$; using $\kappa' \geq 0$ we conclude $M_3(u) \geq f'(u)\kappa(u)$. Moreover, from $f(u) = (u-1)k(u)$ and (3.52c) we conclude $f \in \mathrm{C}^{1,\alpha}([0,u_0])$. Hence, $M_2(u) = \int_0^u f'(u)-f'(\nu)\,\mathrm{d}\nu$ satisfies $|M_2(u)| \leq Cu^{1+\alpha}$. Together we find

$$\eta_2(u) \geq f'(u) - \tfrac{d}{4}C^2u^{2\alpha-1}/\kappa(u) \geq f'(0) - Cu^\alpha - \tfrac{d}{4}C^2\tfrac{u^{2\alpha-1}|\log u|}{(1-u)k(u)} \geq \lambda_2^- \text{ on }]0,u_0]\,.$$

For large u we use the asymptotic formula for $u \to \infty$ given via

$$M_2(u) \approx k(u) + u^2k'(u) \quad\text{and}\quad M_3(u) \approx \tfrac{uk(u)}{2\log u}(k(u)+uk'(u)).$$

Using (3.52d) we find $\eta_2(u) \geq k_2/4 - dk_3/2$ which gives the desired result.

For the last statement note that $M_1(u) = k_0(u{+}1)/2 + \kappa(u) \geq k_0 u/2$, $M_2 \equiv 0$, and $M_3(u) = k_0(\kappa(u) - (u{-}1)\kappa'(u)/2) \geq k_0\kappa(u)/2$. Here, the latter estimate follows from the explicit relation $(u{-}1)\kappa'(u) = (1{-}\frac{\kappa(u)}{k_0 u})\kappa(u)$, cf. [Mie11a, (A.3)]. $\square$

3.3.5 A linear reaction-diffusion system

For $\boldsymbol{u} = (u_1, u_2)$ we consider the system of coupled linear equations

$$\begin{pmatrix} \dot u_1 \\ \dot u_2 \end{pmatrix} = \begin{pmatrix} \delta_1 \Delta u_1 \\ \delta_2 \Delta u_2 \end{pmatrix} + k \begin{pmatrix} u_2{-}u_1 \\ u_1{-}u_2 \end{pmatrix} \quad \text{in } \Omega, \qquad \nabla u_1 \cdot \nu = \nabla u_2 \cdot \nu = 0 \quad \text{on } \partial\Omega, \tag{3.53}$$

which is the gradient flow for the energy $\mathcal{E}$ and the Onsager operator $\mathcal{K}$ given via

$$\begin{aligned} \mathcal{E}(\boldsymbol{u}) &= \int_\Omega u_1(\log u_1{-}1) + u_2(\log u_2{-}1)\,\mathrm{d}x \qquad \text{and} \\ \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} &= \begin{pmatrix} -\mathrm{div}(u_1\delta_1\nabla\xi_1) \\ -\mathrm{div}(u_2\delta_2\nabla\xi_2) \end{pmatrix} + k\Lambda\,(u_1,u_2) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}. \end{aligned} \tag{3.54}$$

The system above describes the diffusion of chemical species A_1 and A_2 which undergo the exchange reaction $A_1 \rightleftharpoons A_2$ with reaction rate k. In particular, observe that the total mass $\mathcal{Q}(u_1,u_2) = \int_\Omega u_1{+}u_2\,\mathrm{d}x$ is conserved along solutions of (3.53), i.e., $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}(u_1,u_2) = 0$. We fix a constant state $\boldsymbol{u}^* = (u_1^*, u_2^*) \in\,]0,\infty[^2$, choose the Sobolev index s as before, and define the spaces

$$\begin{aligned} &H^* = \Big\{\boldsymbol{\xi} \in \mathrm{H}^1(\Omega){\times}\mathrm{H}^1(\Omega) \,:\, \textstyle\int_\Omega \xi_1{+}\xi_2\,\mathrm{d}x = 0\Big\}, \\ &Y = \Big\{\boldsymbol{v} \in \mathrm{H}^{s-2}(\Omega){\times}\mathrm{H}^{s-2}(\Omega) \,:\, \textstyle\int_\Omega v_1{+}v_2\,\mathrm{d}x = 0 \text{ and } \nabla v_1{\cdot}\nu{=}\nabla v_2{\cdot}\nu{=}0\Big\}, \\ &\mathcal{Y} = \{\boldsymbol{u} \in \boldsymbol{u}^*{+}Y \,:\, \inf u_i > 0,\ i=1,2\}, \\ &Z = (H^s(\Omega){\times}\mathrm{H}^s(\Omega)) \cap Y, \quad \mathcal{Z} = \{\boldsymbol{u} \in (\boldsymbol{u}^*{+}Z) \cap \mathcal{Y} \,:\, \nabla(\Delta u_i)\cdot\nu = 0,\ i=1,2\}. \end{aligned}$$

Since $\boldsymbol{u} \mapsto \mathcal{F}(\boldsymbol{u}) = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u})$ is linear we compute

$$\mathrm{D}\mathcal{F}(\boldsymbol{u})[\boldsymbol{v}] = (-\delta_1\Delta v_1 + k(v_1{-}v_2)\,,\ -\delta_2\Delta v_2 + k(v_2{-}v_1))^\mathsf{T}.$$

With the shorthand $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}$ we obtain $\langle \boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = I_1 + I_2$ with

$$\begin{aligned} I_1 &= \int_\Omega \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \cdot \begin{pmatrix} -\delta_1\Delta v_1 + k(v_1{-}v_2) \\ -\delta_2\Delta v_2 + k(v_2{-}v_1) \end{pmatrix} \mathrm{d}x \quad \text{and} \\ I_2 &= -\frac{1}{2}\int_\Omega \big(\delta_1|\nabla\xi_1|^2 {+} k\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\big)\big({-}\delta_1\Delta u_1 {+} k(u_1{-}u_2)\big)\,\mathrm{d}x \\ &\quad - \frac{1}{2}\int_\Omega \big(\delta_2|\nabla\xi_2|^2 {+} k\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1{-}\xi_2)^2\big)\big({-}\delta_2\Delta u_2 {+} k(u_2{-}u_1)\big)\,\mathrm{d}x \end{aligned}$$

Integrating the first term in I_1 by parts twice, using the boundary conditions $\nabla v_i \cdot \nu = \nabla \xi_i \cdot \nu = 0$, and finally substituting $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}$ gives

$$\begin{aligned} I_1 &= \int_\Omega \begin{pmatrix} \delta_1\Delta\xi_1 - k(\xi_1-\xi_2) \\ \delta_2\Delta\xi_2 - k(\xi_2-\xi_1) \end{pmatrix} \cdot \begin{pmatrix} \mathrm{div}(u_1\delta_1\nabla\xi_1) - k\Lambda(\boldsymbol{u})(\xi_1-\xi_2) \\ \mathrm{div}(u_2\delta_2\nabla\xi_2) - k\Lambda(\boldsymbol{u})(\xi_2-\xi_1) \end{pmatrix} \mathrm{d}x \\ &= \int_\Omega \Big[-\delta_1^2 u_1 \nabla\Delta\xi_1 \cdot \nabla\xi_1 - \delta_2^2 u_2 \nabla\Delta\xi_2 \cdot \nabla\xi_2 + k\nabla(\xi_1-\xi_2)\cdot(\delta_1 u_1\nabla\xi_1 - \delta_2 u_2 \nabla\xi_2) \\ &\qquad - k\Lambda(\boldsymbol{u})(\xi_1-\xi_2)(\delta_1\Delta\xi_1 - \delta_2\Delta\xi_2) + 2k^2\Lambda(\boldsymbol{u})(\xi_1-\xi_2)^2 \Big] \mathrm{d}x. \end{aligned}$$

Similarly, we integrate the second term and obtain

$$\begin{aligned} 2I_2 &= \int_\Omega \Big[-\delta_1^2 \nabla u_1 \cdot \nabla\big(|\nabla\xi_1|^2\big) + \big(k\delta_1\Delta u_1 - k^2(u_1-u_2)\big)\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1-\xi_2)^2 \\ &\qquad - \delta_2^2 \nabla u_2 \cdot \nabla\big(|\nabla\xi_2|^2\big) + \big(k\delta_2\Delta u_2 - k^2(u_2-u_1)\big)\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1-\xi_2)^2 \Big] \mathrm{d}x \\ &= \int_\Omega \Big[\delta_1^2 u_1 \Delta\big(|\nabla\xi_1|^2\big) + \big(k\delta_1\Delta u_1 - k^2(u_1-u_2)\big)\partial_{u_1}\Lambda(\boldsymbol{u})(\xi_1-\xi_2)^2 \\ &\qquad + \delta_2^2 u_2 \Delta\big(|\nabla\xi_2|^2\big) + \big(k\delta_2\Delta u_2 - k^2(u_2-u_1)\big)\partial_{u_2}\Lambda(\boldsymbol{u})(\xi_1-\xi_2)^2 \Big] \mathrm{d}x \\ &\quad - \int_{\partial\Omega} \delta_1 u_1 \nabla\big(|\nabla\xi_1|^2\big) + \delta_2 u_2 \nabla\big(|\nabla\xi_2|^2\big)\,\mathrm{d}a. \end{aligned}$$

Thus, using again Bochner's formula and Proposition 3.3.2 we arrive at

$$\begin{aligned} I_1 + I_2 &\geq \int_\Omega \delta_1 u_1 |\mathrm{D}^2\xi_1|^2 + \delta_2^2 u_2 |\mathrm{D}^2\xi_2|^2 + k^2 m(\boldsymbol{u})(\xi_1-\xi_2)^2 + kG(\delta_1,\delta_2,\boldsymbol{u},\boldsymbol{\xi})\,\mathrm{d}x \\ &\text{with } m(\boldsymbol{u}) = 2\Lambda(u_1,u_2) - \tfrac{1}{2}\big(\partial_{u_1}\Lambda(u_1,u_2) - \partial_{u_2}\Lambda(u_1,u_2)\big)(u_1-u_2). \end{aligned}$$

It was shown in [Mie11a, Example 3.5] that $m(\boldsymbol{u}) \geq 2\Lambda(\boldsymbol{u}) \geq 0$ holds.

The main task is to control the mixed terms with prefactor $k\delta_j$ that are collected in the function G. Unfortunately we can estimate these terms only in the case of equal mobilities $\delta_j = \delta > 0$. For $G(\boldsymbol{u},\boldsymbol{\xi}) = \frac{2}{\delta}G(\delta,\delta,\boldsymbol{u},\boldsymbol{\xi})$ some rearrangements yield the identity

$$\begin{aligned} &G(\boldsymbol{u},\boldsymbol{\xi}) = (\xi_1-\xi_2)^2\big(\Delta\Lambda(\boldsymbol{u}) - \mathbb{L}(\boldsymbol{u})\big) - 2\Lambda(\boldsymbol{u})(\xi_1-\xi_2)\Delta(\xi_1-\xi_2) + (u_1+u_2)|\nabla(\xi_1-\xi_2)|^2 \\ &\text{where } \mathbb{L}(\boldsymbol{u}) = \partial_{u_1}^2\Lambda(\boldsymbol{u})|\nabla u_1|^2 + 2\partial_{u_1}\partial_{u_2}\Lambda(\boldsymbol{u})\nabla u_1 \cdot \nabla u_2 + \partial_{u_2}^2\Lambda(\boldsymbol{u})|\nabla u_2|^2. \end{aligned}$$

Since Λ is a concave function, we have $\mathbb{L}(\boldsymbol{u}) \leq 0$. To estimate $\int_\Omega G\,\mathrm{d}x$ we integrate by parts the very first term twice (using $\nabla\boldsymbol{\xi}\cdot\nu = 0$ and $\nabla\Lambda(\boldsymbol{u})\cdot\nu = 0$) and find

$$\int_\Omega G(\boldsymbol{u},\boldsymbol{\xi})\,\mathrm{d}x = \int_\Omega (2\Lambda(\boldsymbol{u}) + u_1 + u_2)|\nabla(\xi_1-\xi_2)|^2 - (\xi_1-\xi_2)^2\mathbb{L}(\boldsymbol{u})\,\mathrm{d}x \geq 0.$$

Hence, we have established the following result.

Theorem 3.3.6 *If Ω is smooth and convex and $\delta_1 = \delta_2 > 0$, then the gradient system* (3.53) *generated by $(\mathcal{E},\mathcal{K})$ from* (3.54) *is geodesically 0-convex.*

3.3.6 Drift-diffusion system in 1D

We consider the one-dimensional version of the drift-reaction-diffusion system (2.7) for electrons and holes in a semiconductor, see Section 2.1.5. We further simplify the system be neglecting the reaction terms $(np-1)$.

To highlight the general structure we treat a system with I nonnegative densities $u_i \in \mathrm{L}^1(\Omega)$ with $\Omega =]0,1[$, where the species have the charge vector $\boldsymbol{q} = (q_i)_{i=1,\dots,I} \in \mathbb{Z}^I$. The system takes the form

$$0 = (\varepsilon\phi_{\boldsymbol{u}}')' + \boldsymbol{q}\cdot\boldsymbol{u}, \qquad \text{in } \Omega, \qquad \phi_{\boldsymbol{u}}(0) = 0 = \phi_{\boldsymbol{u}}'(1); \tag{3.55a}$$

$$\dot u_i = \big[\mu_i(u_i' + u_iV_i' + q_iu_i\phi_{\boldsymbol{u}}')\big]' \qquad \text{in } \Omega; \tag{3.55b}$$

$$0 = \ \mu_i(u_i' + u_iV_i' + q_iu_i\phi_{\boldsymbol{u}}') \qquad \text{for } x \in \{0,1\}, \tag{3.55c}$$

where $'$ is the partial derivative with respect to x. The potentials $\boldsymbol{V} = (V_1,\dots,V_I)$ are smooth functions and contain possible doping terms. The system is the gradient flow for

$$\mathcal{E}(\boldsymbol{u}) = \int_0^1 \sum_{i=1}^I u_i(\log u_i + V_i) + \frac{\varepsilon}{2}|\phi_{\boldsymbol{u}}'|^2 \,\mathrm{d}x \quad \text{and} \quad \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = -\big(\mu_iu_i\xi_i'\big)'_{i=1,\dots,I}. \tag{3.56}$$

Since we have no reaction between the species and no-flux boundary conditions the individual masses $\int_0^1 u_i \,\mathrm{d}x$ are conserved. The electrostatic potential $\phi_{\boldsymbol{u}}$ is a linear function of $\boldsymbol{q}\cdot\boldsymbol{u}$, viz. $\phi_{\boldsymbol{u}} = L\boldsymbol{q}\cdot\boldsymbol{u}$. In the one-dimensional case we have an explicit solution formula:

$$\big(\ \phi = Lg,\ g = \gamma',\ \gamma(1) = 0\ \big) \implies \phi' = -\gamma/\varepsilon. \tag{3.57}$$

The function spaces can be introduced as in the above examples. We only give the calculation of the operator $\mathcal{M}$, where now the quadratic nature of $\mathcal{F}$ due to the terms $u_i\phi_{\boldsymbol{u}}'$ has to be observed. Using the two boundary conditions for $\boldsymbol{\xi} = \mathcal{G}(\boldsymbol{u})\boldsymbol{v}$ we find

$$(\mathrm{D}\mathcal{F}(\boldsymbol{u})^*\boldsymbol{\xi})_i = -\mu_i\xi_i'' + \mu_i(V_i'+q_i\phi_{\boldsymbol{u}}')\xi_i' - q_iLg, \text{ where } g = \textstyle\sum_{j=1}^I \mu_jq_j\big(u_j\xi_j'\big)' = -\boldsymbol{q}\cdot\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi}.$$

Now the quadratic form $\langle\boldsymbol{\xi},\mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = I_1 + I_2$ can be calculated as usual:

$$I_1 = \sum_{i=1}^I \int_0^1 (\mathrm{D}\mathcal{F}(\boldsymbol{u})^*\boldsymbol{\xi})_i(\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi})_i \,\mathrm{d}x = \int_0^1 \sum_{i=1}^I \mu_i^2\Big(-u_i\xi_i'''\xi_i' + u_i\xi_i'\big((V_i'+q_i\phi_{\boldsymbol{u}}')\xi_i'\big)'\Big) + gLg\,\mathrm{d}x,$$

$$I_2 = -\tfrac{1}{2}\sum_{i=1}^I \int_0^1 \mu_i^2(\xi_2')^2\mathcal{F}(\boldsymbol{u})_i \,\mathrm{d}x = \sum_{i=1}^I \mu_i^2 \int_0^1 u_i\big(\xi_i'''\xi_i' + (\xi_i'')^2\big) - \xi_i'\xi_i''u_i(V_i'+q_i\phi_{\boldsymbol{u}}')\,\mathrm{d}x,$$

where we used the boundary conditions $\xi_i' = 0$ on $\partial\Omega$. Combining the two terms and using some cancellation we arrive at

$$\langle\boldsymbol{\xi},\mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \int_0^1 \sum_{i=1}^I \mu_i^2u_i\big((\xi_i'')^2 + V_i''(\xi_i')^2\big) + h_{\boldsymbol{\xi}}\phi_{\boldsymbol{u}}'' + gLg\,\mathrm{d}x \quad \text{with } h_{\boldsymbol{\xi}} = \sum_{i=1}^I \mu_i^2q_iu_i(\xi_i')^2.$$

The first two terms can be estimated in the standard way. For the interaction via $\phi_{\boldsymbol{u}}$ and

L we note that g is such that formula (3.57) can be applied. When assuming additionally that $\varepsilon \equiv \varepsilon_0$ the third and fourth term can be rewritten as

$$Q_{\boldsymbol{u}}(\boldsymbol{\xi}) = h_{\boldsymbol{\xi}} \phi''_{\boldsymbol{u}} + gLg = \frac{1}{\varepsilon_0}\Big(-h_{\boldsymbol{\xi}}\, \boldsymbol{q}\cdot\boldsymbol{u} + \big(\textstyle\sum_{i=1}^{I} \mu_j q_j u_j \xi_j' \big)^2 \Big).$$

There are two cases in which this quadratic form can be estimated from below. First, in the case $I = 1$ we obviously have $Q_{\boldsymbol{u}} \equiv 0$. For $I = 2$ the expression simplifies to

$$Q_{\boldsymbol{u}}(\boldsymbol{\xi}) = -q_1 q_2 u_1 u_2 \big(\mu_1 \xi_1' - \mu_2 \xi_2' \big)^2.$$

Thus, we find $Q_{\boldsymbol{u}} \geq 0$ if $q_1 q_2 \leq 0$, this means that the particles are oppositely charged. Of course, we could add further uncharged particles (i.e. $q_j = 0$), but this is useless as they do not interact with the other particles. We summarize our findings as follows.

Theorem 3.3.7 *Consider the gradient system $(\mathcal{E}, \mathcal{K})$ defined via* (3.55) *and* (3.56) *with constant ε. Assume either $I = 1$ or $I = 2$ and $q_1 q_2 < 0$. If the potentials V_i are λ_i-convex, i.e. $V_i'' \geq \lambda_i$ on Ω, then the gradient system $(\mathcal{E}, \mathcal{K})$ is geodesically λ^*-convex with $\lambda^* = \min\{\mu_i \lambda_i : i = 1, \ldots, I\}$.*

3.3.7 A multi-particle system with cross-diffusion

In several applications one is interested in reaction-diffusion systems with I species, where the microscopic sites are occupied exactly by one species. We refer to [Gri04, BD*10]. On the macroscopic level this means that the density vector $\boldsymbol{u} = (u_1, \ldots, u_I)$ satisfies the pointwise restriction

$$\boldsymbol{u}(x) \cdot \boldsymbol{e} = \sum_{i=1}^{I} u_i(x) = 1 \quad \text{a.e. in } \Omega. \tag{3.58}$$

Moreover, the mobility tensor obeys the Stefan-Maxwell law (see e.g. [Gri04])

$$\mathbb{M}(\boldsymbol{u}) = \mathrm{diag}(\boldsymbol{u}) - \boldsymbol{u} \otimes \boldsymbol{u} = \begin{pmatrix} u_1 - u_1^2 & -u_1 u_2 & \cdots & -u_1 u_I \\ -u_1 u_2 & u_2 - u_2^2 & \cdots & -u_2 u_I \\ \vdots & & \ddots & \vdots \\ -u_1 u_I & -u_2 u_I & \cdots & u_I - u_I^2 \end{pmatrix}. \tag{3.59}$$

Using (3.58) we easily see that $\mathbb{M}$ is positive semidefinite, namely

$$\boldsymbol{a} \cdot \mathbb{M}(\boldsymbol{u})\boldsymbol{a} = \textstyle\sum_{i=1}^{I} u_i a_i^2 - (\boldsymbol{u}\cdot\boldsymbol{a})^2 = \sum_{i=1}^{I} u_i \big(a_i - \boldsymbol{u}\cdot\boldsymbol{a} \big)^2 \geq 0. \tag{3.60}$$

Thus, we consider the energy functional

$$\mathcal{E}(\boldsymbol{u}) = \int_\Omega E(\boldsymbol{u}) + \boldsymbol{u}\cdot\boldsymbol{V}\,\mathrm{d}x, \quad \text{where } E(\boldsymbol{u}) = \sum_{i=1}^{I} u_i(\log u_i - 1) \tag{3.61a}$$

and $\boldsymbol{V} = (V_1, \dots, V_I)$ is a vector of potentials with $\boldsymbol{V} \cdot \boldsymbol{e} \equiv 0$. Thus, $\boldsymbol{V}$ determines the equilibrium state $\boldsymbol{w}$ via $w_i = \mathrm{e}^{-V_i}$. Moreover, the Onsager operator acts now on the vector-valued dual variables $\boldsymbol{\xi} \in H^* = \Big\{ \boldsymbol{\xi} \in \mathrm{H}^1_{\mathrm{av}}(\Omega)^I \ : \ \boldsymbol{\xi} \cdot \boldsymbol{e} \equiv 0 \Big\}$ and takes the form

$$\mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} = \Big(- \operatorname{div}\big(u_i(\nabla \xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\big)\Big)_{i=1,\dots,I} \quad \text{where } \boldsymbol{\Xi}_{\boldsymbol{u}} = \textstyle\sum_{j=1}^I u_j \nabla \xi_j. \tag{3.61b}$$

Taking into account the constraint (3.58) when calculating the differentials we find the nonlinear evolutionary systems

$$\dot{\boldsymbol{u}} = -\mathcal{K}(\boldsymbol{u})\mathrm{D}\mathcal{E}(\boldsymbol{u}) = \Delta \boldsymbol{u} + \Big(\operatorname{div}\big(u_i(\nabla V_i - \boldsymbol{G}_{\boldsymbol{u}})\big)\Big)_{i=1,\dots,I} \quad \text{where } \boldsymbol{G}_{\boldsymbol{u}} = \sum_{j=1}^I u_j \nabla V_j.$$

Here, the diffusion term is linear since $\mathbb{M}(\boldsymbol{u})$ is exactly the inverse of $\mathrm{D}^2 E(\boldsymbol{u})$ (taking the constraint into account). We see that the special choice of $\mathbb{M}$ with negative off-diagonal terms simplifies the diffusion terms, while the drift terms from the potential become more involved. This approach was also used in [Gri04, GaG05], while in [BD*10] the off-diagonal terms are not used.

In particular, the mass of each component is preserved during the flow, namely

$$\int_\Omega \boldsymbol{u}(t,x)\,\mathrm{d}x = \int_\Omega \boldsymbol{u}(0,x)\,\mathrm{d}x = \boldsymbol{m} \in \,]0,\infty[^I \quad \text{with } \boldsymbol{m}\cdot\boldsymbol{e} = \mathrm{vol}(\Omega).$$

In the case $I=2$ the system reduces to a scalar equation for $u \in [0,1]$ via $\boldsymbol{u} = (u, 1{-}u)$ of the form

$$\dot{u} = \Delta u + \operatorname{div}\big((u{-}u^2)\nabla V\big) \quad \text{where } 2V = V_1 = -V_2,$$

which is covered by the analysis treated in Section 3.3.3.

We now restrict to the case $\boldsymbol{V} \equiv 0$ and leave the general case for future research. Our aim is to show that the pure (uncoupled) diffusion is geodesically 0-convex. This statement is nontrivial since the metric $\mathsf{d}_{\mathcal{K}}$ induced by the mobility tensor $\mathbb{M}$ couples the densities in a nontrivial way. However, since $\mathbb{M}(\boldsymbol{u})$ can be estimated from above by $\mathbb{M}_{\mathrm{W}}(\boldsymbol{u}) = \mathrm{diag}(\boldsymbol{u}) \in \mathbb{R}^{I\times I}$ we see that $\mathsf{d}_{\mathcal{K}}$ can be estimated from above by the componentwise Wasserstein distance, i.e.

$$\mathsf{d}_{\mathcal{K}}(\boldsymbol{u}^1, \boldsymbol{u}^2)^2 \leq \mathsf{d}_{\mathrm{W}}(\boldsymbol{u}^1, \boldsymbol{u}^2)^2 = \sum_{i=1}^I \mathsf{d}_{\mathrm{Wass}}(u_i^1, u_i^2)^2.$$

Theorem 3.3.8 *Consider the gradient system $(\mathcal{E}, \mathcal{K})$ defined in* (3.61) *with $\boldsymbol{V} \equiv 0$. Then, $\mathcal{E}$ is geodesically 0-convex with respect to $\mathsf{d}_{\mathcal{K}}$.*

Proof: To estimate the quadratic form $\mathcal{M}$ we assume as usual that Ω is a convex domain with smooth boundary and define the spaces Z, Y, and H as before in the Sobolev space H^s, H^{s-2} and H^1, respectively. Moreover, for the functions $\boldsymbol{u}$, $\boldsymbol{\xi}$, and $\boldsymbol{v} = (v_1, \dots, v_I) = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} \in Y$ we have the following boundary conditions:

$$\text{(a) } \nabla u_i \cdot \nu = 0, \quad \text{(b) } \nabla \xi_i \cdot \nu = 0, \quad \text{(c) } \nabla v_i \cdot \nu = 0. \tag{3.62}$$

Using $\mathcal{F}(\boldsymbol{u}) = -\Delta\boldsymbol{u}$ and $\boldsymbol{\xi} \in \mathcal{G}(\boldsymbol{u})Y$ giving $\boldsymbol{v} = \mathcal{K}(\boldsymbol{u})\boldsymbol{\xi} \in Y$, we have

$$\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \sum_{i=1}^{I}\int_\Omega \xi_i(-\Delta v_i) - \frac{1}{2}(-\Delta u_i)\big(|\nabla\xi_i|^2 - 2\nabla\xi_i\cdot\boldsymbol{\Xi}_{\boldsymbol{u}}\big)\,\mathrm{d}x.$$

Using (c) and (b) we can integrate by parts the first term twice. The second term will be integrated once using (a). After inserting the definition of $\boldsymbol{v}$ we arrive at

$$\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \sum_{i=1}^{I}\int_\Omega \Delta\xi_i\,\mathrm{div}\big(u_i(\nabla\xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\big) - \nabla u_i\cdot\nabla\big(\tfrac{1}{2}|\nabla\xi_i|^2\big) + \nabla u_i\cdot\nabla(\xi_i\cdot\boldsymbol{\Xi}_{\boldsymbol{u}})\big)\,\mathrm{d}x.$$

The first term will now be integrated by part once again by using (b), which also implies $\boldsymbol{\Xi}_{\boldsymbol{u}}\cdot\nu = 0$. Integrating the second term will generate a boundary integral that will be nonnegative by Proposition 3.3.2:

$$\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \int_\Omega \sum_{i=1}^{I} u_i\big(-\nabla\Delta\xi_i\cdot\nabla\xi_i + \Delta(\tfrac{1}{2}|\nabla\xi_i|^2)\big) + \mu(\boldsymbol{u},\boldsymbol{\xi})\,\mathrm{d}x + \beta^1_{\partial\Omega},\ \text{where}$$

$$\mu(\boldsymbol{u},\boldsymbol{\xi}) = \sum_{i,j=1}^{I}\sum_{\alpha,\beta=1}^{d}\big(u_iu_j\xi_{i\alpha\alpha\beta}\xi_{j\beta} + u_{i\beta}u_{j\beta}\xi_{i\alpha}\xi_{j\alpha} + u_{i\beta}u_j\xi_{i\alpha\beta}\xi_{j\alpha} + u_{i\beta}u_j\xi_{i\alpha}\xi_{j\alpha\beta}\big)$$

$$\text{and } \beta^1_{\partial\Omega} = \int_{\partial\Omega}\sum_{i=1}^{I} u_i\mathbb{I}(\nabla_{\|}\xi_i, \nabla_{\|}\xi_i)\,\mathrm{d}a \geq 0.$$

Here, the indices α and β denote partial derivatives with respect to x_α.

The first term in $\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle$ is positive by Bochner's identity. To estimate μ we interchange the summation indices i and j in the fourth term to find that the last two terms can be combined into $(u_iu_j)_\beta\xi_{i\alpha\beta}\xi_{j\alpha}$. Thus, integration by parts, employing Proposition 3.3.2, and exploiting the cancellation of the terms involving $\xi_{i\alpha\alpha\beta}$ gives

$$\textstyle\int_\Omega \mu(\boldsymbol{u},\boldsymbol{\xi})\,\mathrm{d}x = \int_\Omega |\nabla\boldsymbol{u}^T\nabla\boldsymbol{\xi}|^2 - \big|\sum_{i=1}^{I} u_i\mathrm{D}^2\xi_i\big|^2\,\mathrm{d}x + \beta^2_{\partial\Omega},$$

where $\beta^2_{\partial\Omega} = \int_{\partial\Omega} -\mathbb{I}(\boldsymbol{\Xi}_{\boldsymbol{u}}, \boldsymbol{\Xi}_{\boldsymbol{u}})\,\mathrm{d}a$. Here, we used the boundary conditions (b), which give $\boldsymbol{\Xi}_{\boldsymbol{u}}\cdot\nu = 0$ and hence $\sum_{i=1}^{I} u_i\nabla_{\|}\xi_i = \boldsymbol{\Xi}_{\boldsymbol{u}}$ on $\partial\Omega$. The first term in the above integral is nonnegative, while the other two terms are nonpositive. However, they are dominated by the corresponding positive terms obtained earlier, e.g. $\beta^1_{\partial\Omega} + \beta^2_{\partial\Omega} \geq 0$. Using the same rearrangement as in (3.60) we find the final expression

$$\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle = \int_\Omega \sum_{i=1}^{I} u_i|\mathrm{D}^2\xi_i - \boldsymbol{H}|^2 + |\nabla\boldsymbol{u}^T\nabla\boldsymbol{\xi}|^2\,\mathrm{d}x + \int_{\partial\Omega}\sum_{i=1}^{I} u_i\mathbb{I}(\nabla_{\|}\xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}}, \nabla_{\|}\xi_i - \boldsymbol{\Xi}_{\boldsymbol{u}})\,\mathrm{d}a,$$

where $\boldsymbol{H} = \sum_{i=1}^{I} u_i\mathrm{D}^2\xi_i$. Thus, we have established the desired result $\langle\boldsymbol{\xi}, \mathcal{M}(\boldsymbol{u})\boldsymbol{\xi}\rangle \geq 0$.

□

4 Multiscale limits

In this chapter we discuss the derivation of limit systems from multiscale problems that exhibit a gradient structure $(X, \mathcal{E}, \Psi)$. In particular, we consider here the following two model problems: the derivation of bulk/surface coupling for the Allen-Cahn equation and of an effective interface condition for a one-dimensional diffusion equation.

Here, the crucial point is that in the analysis of these problems we only rely on the gradient structure of the systems. More precisely, we consider families of driving functionals $\mathcal{E}_\varepsilon$ and dissipation potentials Ψ_ε, where the parameter ε describes the different scales in the system. Assuming that $\mathcal{E}_\varepsilon$ and Ψ_ε converge in a variational sense to limit functionals $\mathcal{E}_0$ and Ψ_0 for $\varepsilon \to 0$ the obvious question is: Do solutions with respect to $\mathcal{E}_\varepsilon$ and Ψ_ε converge in some sense to solutions of the limit system?

Our approach to this question is based upon the De Giorgi principle which characterizes solutions as curves of maximal slope (see [AGS05]) and is written as the energy-dissipation equation

$$\forall\, t \in [0,T]: \quad \mathcal{E}_\varepsilon(u_\varepsilon(t)) + \int_0^t \left[\Psi_\varepsilon(u_\varepsilon; \dot{u}_\varepsilon) + \Psi_\varepsilon^*(u_\varepsilon; -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))\right] \mathrm{d}s = \mathcal{E}_\varepsilon(u_\varepsilon(0)).$$

This formulation opens the door for the application of notions of variational convergence such as Mosco and Γ-convergence. We refer to the monographs by Attouch [Att84], Dal Maso [Dal93], and Braides [Bra02]) for comprehensive survey of the theory of variational convergence of functionals.

In Section 4.1 we study the derivation of an Allen-Cahn equation with bulk/surface coupling. This means that the evolution in the bulk, described by the Allen-Cahn equation, is coupled to an evolutionary system defined on the boundary. Such problems can be found e.g. in the theory of spinodal decomposition of binary mixtures (see e.g. [KE*01]). Starting point of our discussion in Section 4.1.1 is a fixed domain (bulk) which is surrounded by boundary layer of thickness ε. In the bulk and the boundary layer the system's evolution is given by an Allen-Cahn equations, respectively, whose gradient structure was discussed in Section 2.1.1. Bulk and boundary layer system are coupled by natural continuity and transmission conditions at the common boundary. Additionally, we assume that the coefficients in the boundary layer satisfy certain scalings in terms of the parameter ε, see (4.3). After a rescaling of the boundary layer, which is introduced in 4.1.2, we use the Mosco convergence of the Allen-Cahn energies and the dissipation potentials to pass to the limit in the De Giorgi principle in Section 4.1.3. Here, we follow the ideas of Sandier & Serfaty [SaS04] (see also [BFG06, KMM06, Kur07]), where an approach to prove the convergence of gradient flows for Γ-converging energy functionals was presented and applied to derive the limiting dynamics of vortices for the heat flow of the Ginzburg-Landau energy. Moreover, we discuss the possibility of passing to the limit in the evolutionary variational inequality (EVI_λ) when the energy functionals are λ-convex. Here, we exploit the simple and flat geometry induced by the L^2-metric.

In the second example in Section 4.2 we discuss the derivation of an interface condition for the one-dimensional (linear) diffusion equation. Starting point of this derivation is a bulk system with a small interface layer of width ε around the origin. We assume that the diffusion coefficient in the interface layer is of order ε. The evolution of this system is given as the gradient flow of the logarithmic free energy and a Wasserstein-type Onsager operator with spatially dependent diffusion coefficient. Using the De Giorgi principle we show in Section 4.2.2 that solutions converge (up to subsequences) to a solution of the bulk/interface system

$$\dot{u} = \rho u'' \quad \text{in } \Omega\backslash\{0\}, \qquad \rho u'_+ = \rho u'_- = k(u_+ - u_-) \quad \text{at } \{0\},$$

where $_+$ and $_-$ denote the limits from the right and left, respectively. In particular, the proof of the convergence uses a rescaling of the interface layer and follows the ideas in [AM*12]. There, a similar limit was discussed, namely the passage from diffusion in a one-dimensional Fokker-Planck equation to (linear) reaction. The basic idea here is to exploit the compactness properties of the solutions and show liminf estimates for the energies and the dissipations. Then, the lower bounds in the liminf estimate can be characterized as the limit gradient system corresponding to the equation above (see Theorem 4.2.5). In Section 4.2.3 we comment on the possibility of using the methods of Chapter 3, i.e., the geodesic λ-convexity of the driving functional $\mathcal{E}$ and the (EVI_λ) formulation, to derive the limit system. It turns out that we are not able to exploit this property as we have no uniform geodesic λ_ε-convexity, i.e., $\lambda_\varepsilon \to -\infty$ when ε goes to zero.

4.1 Bulk/surface evolution for the Allen-Cahn equation

In the recent years there has been a growing interest in the coupling of bulk and surface processes. One important example is the theory of spinodal decomposition of binary mixtures where dynamic boundary conditions are used to model the effective short-range interaction between the two mixture components and the wall (i.e., the boundary), see e.g. [Kra95, PuF97] and the references therein. Moreover, we refer to [KE*01, RaZ01, MiZ05, CFP06, FRG*06, CGM08, GGM08, SpW10] for an (incomplete) list of articles related to the mathematical analysis of dynamic boundary conditions for various evolutionary systems including the heat equation, the iso- and non-isothermal Allen-Cahn equations, the Cahn-Hilliard equation and the Caginalp system. In addition, we point out to the book [Tai09] for the connection to Feller semigroups and Markov processes.

In this section we present the results of [Lie12] where the question was discussed whether such dynamic boundary conditions can be obtained as a limit of a family of bulk systems in the case of the Allen-Cahn equation. The relevance of this question lies in the identification of the relevant scalings in the approximating boundary layer system in order to obtain more information about the structure of the limit systems.

A similar question was answered in a non-rigorous fashion in [CoR90] for the linear heat equation. Moreover, we refer to [ScT10] for the related problem of deriving models for conductive thin sheets where the method of asymptotic expansion was used.

4.1.1 Setting of the model

Let us consider an open and bounded domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, with a C^2-boundary denoted by $\Gamma \stackrel{\text{def}}{=} \partial\Omega$. For a sufficiently small $\varepsilon > 0$ we shall introduce the domain Ω_ε defined by

$$\Omega_\varepsilon \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^d \,:\, \inf_{y\in\Omega} |x-y| < \varepsilon \right\}.$$

We call the set $\Sigma_\varepsilon \stackrel{\text{def}}{=} \Omega_\varepsilon \setminus \overline{\Omega}$ the boundary layer of Ω. Obviously, we have the convergence $\Omega_\varepsilon \to \Omega$ for $\varepsilon \to 0$ with respect to the Hausdorff distance.

Denoting a finite time horizon by $T > 0$ we consider the following system of Allen-Cahn equations:

$$\begin{aligned} \tau_\mathrm{b}\dot{u}_\varepsilon &= A_\mathrm{b}\Delta u_\varepsilon - W_\mathrm{b}'(u_\varepsilon) && \text{in } [0,T]\times\Omega, \\ \tau_\varepsilon\dot{u}_\varepsilon &= A_\varepsilon\Delta u_\varepsilon - \frac{1}{\varepsilon}W_\mathrm{s}'(u_\varepsilon) && \text{in } [0,T]\times\Sigma_\varepsilon, \end{aligned} \tag{4.1}$$

where $\tau_\mathrm{b},\ \tau_\varepsilon > 0$ denote the relaxation times, $A_\mathrm{b},\ A_\varepsilon$ the diffusion coefficients, and W_b', W_s' are the derivatives of potentials $W_\mathrm{b}, W_\mathrm{s} \in \mathrm{C}^1(\mathbb{R})$ in the bulk and in the boundary layer, respectively. The system above is subjected to the following natural boundary and transmission conditions at $\partial\Omega_\varepsilon$ and at the common boundary Γ

$$\begin{aligned} A_\varepsilon \tfrac{\partial u_\varepsilon}{\partial\nu} &= 0 && \text{on } [0,T]\times\partial\Omega_\varepsilon, \\ A_\mathrm{b}\tfrac{\partial u_\varepsilon}{\partial\nu} &= A_\varepsilon\tfrac{\partial u_\varepsilon}{\partial\nu} && \text{on } [0,T]\times\Gamma, \\ [\![u_\varepsilon]\!] &= 0 && \text{on } [0,T]\times\Gamma, \end{aligned} \tag{4.2}$$

where ν denotes the outer unit normal on Γ and $\partial\Omega_\varepsilon$ and $[\![\,\cdot\,]\!]$ denotes the jump across the common boundary Γ. The system is completed by imposing the initial condition $u_\varepsilon(0) = u_\varepsilon^0$.

In order to derive nontrivial boundary conditions when ε goes to 0 we let the relaxation time τ_ε and the diffusion coefficient A_ε depend on ε in Σ_ε. In particular, we assume that in Σ_ε the coefficients satisfy the scalings

$$\tau_\varepsilon = \varepsilon^{-\alpha}\tau_\mathrm{s} \qquad \text{and} \qquad A_\varepsilon = \varepsilon^{-\beta}A_\mathrm{s} \tag{4.3}$$

for given $\tau_\mathrm{s}, A_\mathrm{s} > 0$ and $\alpha \in \mathbb{R}$, $\beta \in \,]-1,\infty[$. This amounts to different length and time scales in the bulk and in the boundary layer.

The nonlinearities W_b and W_s are at least of quadratic growth and satisfy the growth conditions

$$|W_{\mathrm{b/s}}'(u)| \leq C(1+|u|^p) \ \text{ with } \ p \in [1,q[\ \text{ and } \ q = \begin{cases} \infty & d=2, \\ \frac{d+2}{d-2} & d \geq 3. \end{cases} \tag{4.4}$$

These are the same growth conditions imposed in [SpW10] for the bulk potential W_b, while we have a stronger growth condition for the boundary potential since we are in the full d-dimensional domain Σ_ε in contrast to the $(d-1)$-dimensional boundary Γ in [SpW10].

We show that solutions of the system above converge in a certain sense to a solution of a limit system which consists of the bulk equation in Ω in (4.1) coupled to an equation posed on the boundary Γ. Obviously, the form of the latter equation will heavily depend on the choices for the scaling exponents α and β.

To put the above system in an abstract framework we introduce the function spaces

$$V_\varepsilon \stackrel{\mathrm{def}}{=} \mathrm{H}^1(\Omega_\varepsilon) \qquad \text{and} \qquad H_\varepsilon \stackrel{\mathrm{def}}{=} \mathrm{L}^2(\Omega_\varepsilon).$$

Then, the weak formulation of the system (4.1) reads: Find $u_\varepsilon \in \mathrm{L}^2(0,T;V_\varepsilon)$ with $\dot{u}_\varepsilon \in \mathrm{L}^2(0,T;H_\varepsilon)$ such that for all $\varphi \in V_\varepsilon$ and almost all $t \in [0,T]$ we have

$$0 = \int_{\Omega_\varepsilon} \Big[\mathbb{G}_\varepsilon(x)\dot{u}_\varepsilon(t)\varphi + \mathbb{A}_\varepsilon(x)\nabla u_\varepsilon(t)\cdot\nabla\varphi + \mathbb{W}_\varepsilon'(x,u_\varepsilon(t))\varphi\Big]\,\mathrm{d}x, \tag{4.5}$$

where we use the notation

$$(\mathbb{G}_\varepsilon(x),\mathbb{A}_\varepsilon(x)) = \begin{cases}(\tau_\mathrm{b},A_\mathrm{b}) & \text{for } x\in\Omega,\\ (\tau_\varepsilon,A_\varepsilon) & \text{for } x\in\Sigma_\varepsilon,\end{cases} \qquad \mathbb{W}_\varepsilon(x,\cdot) = \begin{cases} W_\mathrm{b}(\cdot) & \text{for } x\in\Omega,\\ \frac{1}{\varepsilon}W_\mathrm{s}(\cdot) & \text{for } x\in\Sigma_\varepsilon.\end{cases}$$

The existence of solutions of (4.5) follows from standard arguments, see e.g. [Rou05, SpW10].

Theorem 4.1.1 (Existence of solutions) *For fixed $\varepsilon>0$ let $u_\varepsilon^0 \in V_\varepsilon$ be given. Moreover, assume that the growth condition* (4.4) *holds. Then, there exists a solution $u_\varepsilon \in \mathrm{H}^1(0,T;\mathrm{L}^2(\Omega_\varepsilon)) \cap \mathrm{L}^\infty(0,T;\mathrm{H}^1(\Omega_\varepsilon))$ of the system* (4.1).

In Subsection 2.1.1 we noted that equation (4.1) is the L^2-gradient flow of the Allen-Cahn functional $\mathcal{E}_\varepsilon : V_\varepsilon \to \mathbb{R}$ defined by

$$\mathcal{E}_\varepsilon(u) = \int_{\Omega_\varepsilon} \Big[\frac{\mathbb{A}_\varepsilon(x)}{2}|\nabla u|^2 + \mathbb{W}_\varepsilon(x,u)\Big]\,\mathrm{d}x.$$

More precisely, by defining the multiplication operator $\mathcal{G}_\varepsilon : H_\varepsilon \to H_\varepsilon^*$ via $\langle \mathcal{G}_\varepsilon \dot{u}, \dot{v}\rangle = \int_{\Omega_\varepsilon} \mathbb{G}_\varepsilon(x)\dot{u}\dot{v}\,\mathrm{d}x$ the equation in (4.5) can then be written in the form

$$\mathcal{G}_\varepsilon \dot{u}_\varepsilon(t) = -\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t)), \tag{4.6}$$

with $\mathrm{D}\mathcal{E}_\varepsilon(u) \in V_\varepsilon^*$ denoting the Gâteaux derivative of $\mathcal{E}_\varepsilon$ which is well-defined due to (4.4). Defining the inverse operator $\mathcal{K}_\varepsilon {=} \mathcal{G}_\varepsilon^{-1}{:}H_\varepsilon^* {\to} H_\varepsilon$ we can rewrite (4.6) as

$$\dot{u}_\varepsilon(t) = -\mathcal{K}_\varepsilon \mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon(t)) =: -\nabla_{\mathcal{G}_\varepsilon}\mathcal{E}_\varepsilon(u_\varepsilon(t)), \tag{4.7}$$

where $\nabla_{\mathcal{G}_\varepsilon}\mathcal{E}$ denotes the gradient of $\mathcal{E}_\varepsilon$ with respect to the metric tensor $\mathcal{G}_\varepsilon$. Note that we have $\langle \xi, \mathcal{K}_\varepsilon \eta\rangle = \int_{\Omega_\varepsilon} \mathbb{G}_\varepsilon(x)^{-1}\xi\eta\,\mathrm{d}x$. The operator $\mathcal{G}_\varepsilon$ defines the quadratic dissipation potential $\Psi_\varepsilon(\dot u) = \frac12\langle \mathcal{G}_\varepsilon \dot u, \dot u\rangle$ whose Legendre transform is in turn given by $\mathcal{K}_\varepsilon$, i.e., we have $\Psi^*_\varepsilon(\xi) = \frac12\langle \xi, \mathcal{K}_\varepsilon \xi\rangle$, where ξ is the dual variable to $\dot u$.

With this we can write the evolution of the system equivalently using the De Giorgi principle

$$\mathcal{E}_\varepsilon(u_\varepsilon(0)) - \mathcal{E}_\varepsilon(u_\varepsilon(t)) = \int_0^t \Big[\Psi_\varepsilon(\dot u_\varepsilon) + \Psi^*_\varepsilon(-\mathrm{D}\mathcal{E}_\varepsilon(u_\varepsilon))\Big]\,\mathrm{d}s. \tag{4.8}$$

We study the behavior of the solutions u_ε when $\varepsilon \to 0$. In this case the boundary layer Σ_ε shrinks to Γ, and we show that the "limit" of the sequence $u_\varepsilon|_{\Sigma_\varepsilon}$ describes the evolution on Γ.

4.1.2 Transformation of the problem

In order to provide a notion of convergence of the solutions u_ε we transform the variable domain Ω_ε to a fixed domain.

Due to the smoothness of the boundary Γ we can characterize a point $x \in \Sigma_\varepsilon$ for sufficiently small ε in the following way: There exist unique $y \in \Gamma$ and $\vartheta \in\,]0,\varepsilon[$ such that $x = y + \vartheta\nu(y)$ (see e.g. [Wł87, Chapter 2]). Hence, we introduce the change of coordinates in Σ_ε (see Figure 4.1)

$$\begin{aligned} x_\varepsilon(y,\theta) &\overset{\text{def}}{=} y + \varepsilon\theta\nu(y), && (y,\theta) \in \Gamma\times]0,1[,\\ (y_\varepsilon(x),\theta_\varepsilon(x)) &\overset{\text{def}}{=} (\mathsf{P}_\varepsilon(x), \mathsf{d}_\varepsilon(x)/\varepsilon), && x \in \Omega_\varepsilon, \end{aligned}$$

where P_ε and d_ε denote the projection from Σ_ε on Γ and the distance to Γ, respectively.

With this change of coordinates we define $\Sigma \overset{\text{def}}{=} \Gamma\times]0,1[$ and for a function $u : \Sigma_\varepsilon \to \mathbb{R}$ we set $U = u \circ x_\varepsilon : \Sigma \to \mathbb{R}$. Since the boundary Γ is of class C^2 we have that the outer unit normal satisfies $\nu \in \mathrm{C}^1(\Gamma;\mathbb{R}^d)$. Therefore, if $u \in \mathrm{H}^1(\Sigma_\varepsilon)$ we have $U \in \mathrm{H}^1(\Sigma)$. More precisely, u and U satisfy

$$\begin{pmatrix} \nabla_\Gamma U \\ \partial_\theta U \end{pmatrix} = \begin{pmatrix} \mathbb{P}(y) - \varepsilon\theta\,\mathbb{S}(y) \\ \varepsilon\,\nu(y)^\top \end{pmatrix} \nabla u, \quad \text{and} \quad \nabla u = \Big(\mathbb{Q}_\varepsilon(x) \Big| \frac{1}{\varepsilon}\nu(\mathsf{P}_\varepsilon(x)) \Big) \begin{pmatrix} \nabla_\Gamma U \\ \partial_\theta U \end{pmatrix},$$

where $\nabla_\Gamma U \in \mathcal{T}(\Gamma)$ denotes the tangential gradient of U on Γ, $\mathbb{P}(y)$ is the projection onto the tangential space at $y \in \Gamma$, $\mathbb{S} = -\mathrm{D}\nu$ is the so-called shape operator (see e.g. [dCa76]) and $\mathbb{Q}_\varepsilon$ is such that $\mathbb{Q}_\varepsilon(\mathbb{P} - \varepsilon\theta\mathbb{S}) = \mathbb{P}$.

The tangential gradient $\nabla_\Gamma U$ on Γ can be characterized in the following way (see [SaV97, dCa76]): For $U : \Gamma \to \mathbb{R}$ we denote by $\widetilde U$ a smooth extension of U to $\mathbb{R}^d$, then $\nabla_\Gamma U(y) = \mathbb{P}(y)[\nabla \widetilde U]$. It is easy to check that this definition is well-defined and independent of the extension $\widetilde U$, moreover, we have that $\mathbb{P} = I - \nu \otimes \nu$. Similarly, the divergence on Γ for vector fields $\boldsymbol{X}$ on Γ can be defined as

$$\mathrm{div}_\Gamma \boldsymbol{X} = \mathrm{div}\widetilde{\boldsymbol{X}} - \nabla(\widetilde{\boldsymbol{X}}\cdot\nu)\nu,$$

where $\widetilde{\boldsymbol{X}}$ denotes analogously a smooth extension of $\boldsymbol{X}$. The Laplace-Beltrami operator Δ_Γ on Γ is then simply given as $\Delta_\Gamma U = \mathrm{div}_\Gamma(\nabla_\Gamma U)$. For a vector field $\boldsymbol{X} \in \mathrm{L}^2(\Gamma;\mathcal{T}(\Gamma))$

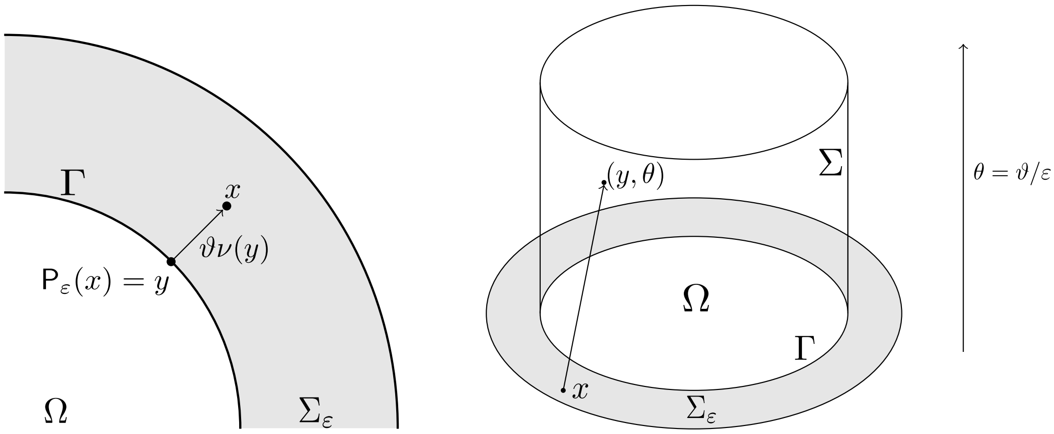

Figure 4.1: Transformation of the boundary layer.

such that $\mathrm{div}_\Gamma \boldsymbol{X} \in \mathrm{L}^2(\Gamma)$ and $U \in \mathrm{H}^1(\Gamma)$ we have Green's formula

$$-\int_\Gamma \nabla_\Gamma U \cdot \boldsymbol{X}\,\mathrm{d}\Gamma = \int_\Gamma U \mathrm{div}_\Gamma \boldsymbol{X}\,\mathrm{d}\Gamma.$$

We leave the bulk domain Ω untransformed. Thus, we introduce the spaces for the bulk and surface component $u : \Omega \to \mathbb{R}$ and $U : \Sigma \to \mathbb{R}$ as

$$\mathcal{V} \overset{\text{def}}{=} \Big\{(u,U) \in \mathrm{H}^1(\Omega)\times\mathrm{H}^1(\Sigma) \;:\; u|_\Gamma = U|_{\{\theta=0\}}\Big\}, \qquad \mathcal{H} \overset{\text{def}}{=} \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Sigma).$$

Here, the measure on Σ is given by $\mathrm{d}\mu = \mathrm{d}\Gamma \otimes \mathrm{d}\lambda^1$, i.e., the product of the surface measure on Γ and the one-dimensional Lebesgue measure on $]0,1[$. The space $\mathrm{H}^1(\Sigma)$ is defined in the usual way, i.e., the closure of $\mathrm{C}^1(\Sigma)$ with respect to the norm $\|\cdot\|_{\mathrm{H}^1(\Sigma)}$, where

$$\|U\|^2_{\mathrm{H}^1(\Sigma)} = \int_\Sigma \Big[|U|^2 + |\nabla_\Gamma U|^2 + |\partial_\theta U|^2\Big]\,\mathrm{d}\mu.$$

Now, substituting the above transformations in $\mathcal{E}_\varepsilon$ we arrive at the transformed energy functional $\overline{\mathcal{E}}_\varepsilon : \mathcal{V} \to \mathbb{R}$ for $\boldsymbol{u} = (u,U)$ defined by

$$\begin{aligned}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}) = &\int_\Omega \left[\frac{A_\mathrm{b}}{2}|\nabla u|^2 + W_\mathrm{b}(u)\right]\mathrm{d}x \\ &+ \int_\Sigma \left[\frac{A_\varepsilon}{2}\left(\nabla_\Gamma U \cdot \mathbb{B}_\varepsilon(y,\theta)\nabla_\Gamma U + \frac{1}{\varepsilon^2}|\partial_\theta U|^2\right) + \frac{1}{\varepsilon}W_\mathrm{s}(U)\right]\mathbb{J}_\varepsilon(y,\theta)\,\mathrm{d}\mu,\end{aligned}$$

where $\mathbb{B}_\varepsilon = \mathbb{Q}_\varepsilon^\top \mathbb{Q}_\varepsilon$ and $\mathbb{J}_\varepsilon$ describes the change of volume due to the transformation. Additionally, the transformed dissipation potential $\overline{\Psi}_\varepsilon : \mathcal{H} \to [0,\infty[$ reads

$$\overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}) = \int_\Omega \frac{\tau_\mathrm{b}}{2}|\dot{u}|^2\,\mathrm{d}x + \int_\Sigma \frac{\tau_\varepsilon}{2}|\dot{U}|^2 \mathbb{J}_\varepsilon(y,\theta)\,\mathrm{d}\mu.$$

We denote by $\overline{\mathcal{G}}_\varepsilon : \mathcal{H} \to \mathcal{H}^*$ the associated operator, i.e., $\overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}) = \frac{1}{2}\langle \overline{\mathcal{G}}_\varepsilon \dot{\boldsymbol{u}}, \dot{\boldsymbol{u}}\rangle$. The inverse operator $\overline{\mathcal{K}}_\varepsilon = \overline{\mathcal{G}}_\varepsilon^{-1} : \mathcal{H}^* \to \mathcal{H}$ gives the dual dissipation potential $\overline{\Psi}_\varepsilon^*$, more precisely, for a dual variable $\boldsymbol{\xi} = (\xi, \Xi)$ it reads

$$\overline{\Psi}_\varepsilon^*(\boldsymbol{\xi}) = \int_\Omega \frac{\tau_{\mathrm{b}}^{-1}}{2} |\xi|^2 \,\mathrm{d}x + \int_\Sigma \frac{\tau_\varepsilon^{-1}}{2\mathbb{J}_\varepsilon(y,\theta)} |\Xi|^2 \,\mathrm{d}\mu.$$

Note that although we have that $\mathcal{E}_\varepsilon(u) = \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u})$ and $\Psi_\varepsilon(\dot{u}) = \overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}})$ it holds that $\Psi_\varepsilon^*(\xi) \neq \overline{\Psi}_\varepsilon^*(\boldsymbol{\xi})$. This is due to the fact that the Legendre transform $\overline{\Psi}_\varepsilon^*$ is calculated in the space $\mathcal{H}$ whose norm and scalar product are not inherited from H_ε. For the same reasons we have that $\mathrm{D}\mathcal{E}_\varepsilon(u) \neq \mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u})$. However, we have the identity $\Psi_\varepsilon(-\mathrm{D}\mathcal{E}_\varepsilon(u)) = \overline{\Psi}_\varepsilon(-\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}))$. In particular, the energy-dissipation equation (4.8) is equivalent to

$$\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(t)) + \int_0^t \Big[\overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}_\varepsilon) + \overline{\Psi}_\varepsilon^*(-\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon))\Big] \,\mathrm{d}s = \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(0)). \tag{4.9}$$

The following lemma is concerned with the convergences of the geometrical quantities $\mathbb{B}_\varepsilon$ and $\mathbb{J}_\varepsilon$.

Lemma 4.1.2 *$\mathbb{B}_\varepsilon$ converges uniformly in Σ to $\mathbb{I}$, with $\mathbb{I}$ denoting the identity in the tangent bundle of Γ, and $\mathbb{J}_\varepsilon/\varepsilon$ converges uniformly in Σ to 1.*

The easiest way to see that the convergence is indeed as stated, is to switch to local coordinates and calculate $\mathbb{B}_\varepsilon$ and $\mathbb{J}_\varepsilon$ explicitly in terms of the covariant and contravariant basis vectors (see [Cia00, Section 1.2] for a related problem in the theory of elastic shells).

4.1.3 Convergence of the system

Our result is formulated abstractly in terms of Mosco convergence of $\overline{\mathcal{E}}_\varepsilon$ towards a limit $\overline{\mathcal{E}}_0$ and of $\overline{\Psi}_\varepsilon$ towards $\overline{\Psi}_0$. For functionals $\mathcal{F}_n$, defined on a Banach space $\mathcal{Q}$, the definition of Mosco convergence is as follows:

$$\mathcal{F}_n \xrightarrow{\mathrm{M}} \mathcal{F} \;\Leftrightarrow\; \begin{cases} \text{(i) Liminf estimate for weakly converging sequences:} \\ \quad \boldsymbol{q}_n \rightharpoonup \boldsymbol{q} \quad \Longrightarrow \quad \mathcal{F}(\boldsymbol{q}_n) \le \liminf\limits_{n\to\infty} \mathcal{F}_n(\boldsymbol{q}), \\ \text{(ii) Existence of strongly converging recovery sequences:} \\ \quad \forall\, \widehat{\boldsymbol{q}} \in \mathcal{Q}\ \exists\, (\widehat{\boldsymbol{q}}_n)_n :\ \widehat{\boldsymbol{q}}_n \to \widehat{\boldsymbol{q}} \text{ and } \mathcal{F}(\widehat{\boldsymbol{q}}) \ge \limsup\limits_{n\to\infty} \mathcal{F}_n(\widehat{\boldsymbol{q}}_n). \end{cases}$$

Hence, Mosco convergence is nothing but Γ-convergence in the weak *and* in the strong topology of the Banach space $\mathcal{Q}$.

Since it is essential to choose the right topology for computing the Γ- or Mosco limits, the first step in our convergence proof is to derive a priori estimates for the solutions $(u_\varepsilon, U_\varepsilon)$. This is addressed in the following lemma.

Lemma 4.1.3 (A priori estimate) *Let $\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(0)) \le C < \infty$ and define $\Omega_T = \Omega \times \,]0,T[$ and $\Sigma_T = \Sigma \times \,]0,T[$. Then, there exist constants $C_1, C_2, C_3, C_4 > 0$, independent of ε,* *such that*

$$
\begin{aligned}
\|\dot{u}_\varepsilon\|^2_{\mathrm{L}^2(\Omega_T)} + \varepsilon^{1-\alpha}\|\dot{U}_\varepsilon\|^2_{\mathrm{L}^2(\Sigma_T)} &\le C_1,\\
\|\mathrm{D}_u\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)\|^2_{\mathrm{L}^2(\Omega_T)} + \frac{1}{\varepsilon^{1-\alpha}}\|\mathrm{D}_U\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)\|^2_{\mathrm{L}^2(\Sigma_T)} &\le C_2,
\end{aligned}
\tag{4.10a}
$$

and

$$
\begin{aligned}
\|\nabla u_\varepsilon(t)\|^2_{\mathrm{L}^2(\Omega)} + \|u_\varepsilon(t)\|^2_{\mathrm{L}^2(\Omega)} + \|U_\varepsilon(t)\|^2_{\mathrm{L}^2(\Sigma)} &\le C_3,\\
\varepsilon^{1-\beta}\|\nabla_\Gamma U_\varepsilon(t)\|^2_{\mathrm{L}^2(\Sigma)} + \frac{1}{\varepsilon^{1+\beta}}\|\partial_\theta U_\varepsilon(t)\|^2_{\mathrm{L}^2(\Sigma)} &\le C_4,
\end{aligned}
\tag{4.10b}
$$

for all $t \in [0,T]$.

Proof: The estimates in (4.10) are a direct consequence of the energy balance (4.9). We remind that the relaxation time and the diffusion coefficient are given by $\tau_\varepsilon = \tau_\mathrm{s}\varepsilon^{-\alpha}$, $A_\varepsilon = A_\mathrm{s}\varepsilon^{-\beta}$. The energy functional satisfies the estimate

$$
\begin{aligned}
\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \ge C(\|\nabla u_\varepsilon\|^2_{\mathrm{L}^2(\Omega)} + \|u_\varepsilon\|^2_{\mathrm{L}^2(\Omega)} + \|U_\varepsilon\|^2_{\mathrm{L}^2(\Omega)}&\\
+ \varepsilon^{1-\beta}\|\nabla_\Gamma U_\varepsilon\|^2_{\mathrm{L}^2(\Sigma)} + \varepsilon^{-(\beta+1)}\|\partial_\theta U_\varepsilon\|^2_{\mathrm{L}^2(\Sigma)}) - c,&
\end{aligned}
$$

where we have used the quadratic growth of the nonlinearities W_b and W_s as well as Lemma 4.1.2. The dissipation potential satisfies

$$
\begin{aligned}
\overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}_\varepsilon) &\ge C(\|\dot{u}_\varepsilon\|^2_{\mathrm{L}^2(\Omega)} + \varepsilon^{1-\alpha}\|\dot{U}_\varepsilon\|^2_{\mathrm{L}^2(\Sigma)}),\\
\overline{\Psi}^*_\varepsilon(-\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)) &\ge C(\|\mathrm{D}_u\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)\|^2_{\mathrm{L}^2(\Omega)} + \varepsilon^{\alpha-1}\|\mathrm{D}_U\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)\|^2_{\mathrm{L}^2(\Sigma)}).
\end{aligned}
$$

By assumption the lefthand-side in the energy balance (4.9) is bounded, thus we arrive at (4.10). ☐

The estimates in Lemma 4.1.3 show that the critical scaling for the relaxation time $\tau_\varepsilon = \varepsilon^{-\alpha}\tau_\mathrm{s}$ is $\alpha=1$. For $\alpha<1$ we expect the time derivatives in Σ to blow up while the thermodynamically conjugated driving forces tend to 0 in the limit. This means that we have a much faster timescale in the boundary layer, such that in the limit the system is always in equilibrium on the boundary. Conversely, $\alpha>1$ amounts to a slower timescale in the boundary layer with no evolution. In contrast to these degenerate cases $\alpha=1$ results in a nontrivial dynamic boundary condition as in [SpW10].

In addition, we find the characteristic values $\beta \in \{-1,+1\}$ for the scalings of the diffusion coefficient $A_\varepsilon = \varepsilon^{-\beta}A_\mathrm{s}$ in the boundary layer. For $\beta>1$ all derivatives have to vanish such that U is constant (in every connected component of Σ). However, it is not fixed and may evolve in time, we refer to this as the fast diffusion case. Conversely, for $\beta<1$ we expect the tangential derivatives to blow up in the boundary layer (no diffusion case). For $\beta=1$ we expect genuine surface diffusion.

The crucial point is that in all of the cases above the derivative with respect to θ has to vanish. Hence, in the limit the surface variable U is given only by its trace on Γ which allows for the reduction to surface evolution, see Section 4.1.4 for the final discussion.

Lemma 4.1.3 shows that we can extract a (not relabeled) subsequence $\boldsymbol{u}_\varepsilon = (u_\varepsilon, U_\varepsilon)$

such that for the bulk variable u_ε we have the convergence

$$\begin{aligned} u_\varepsilon &\overset{*}{\rightharpoonup} u && \text{in } \mathrm{L}^\infty(0,T;\mathrm{H}^1(\Omega)),\\ \dot{u}_\varepsilon &\rightharpoonup \dot{u} && \text{in } \mathrm{L}^2(\Omega_T). \end{aligned} \tag{4.11}$$

The second estimate in (4.10a) shows (by eventually extracting another subsequence) that we have the convergence $\mathrm{D}\mathcal{E}_{\mathrm{b}}(u_\varepsilon) \rightharpoonup \xi$ in $\mathrm{L}^2(\Omega_T)$, where $\mathcal{E}_{\mathrm{b}}(u)$ denotes the bulk energy part. However, due to (4.11) we can argue that $u_\varepsilon \to u$ in $\mathrm{L}^q(\Omega_T)$ with $1 \le q < \infty$ for $d = 2$ and $1 \le q < 2d/(d-2)$ for $d \ge 3$. In particular, considering an almost everywhere converging subsequence and using the growth condition (4.4) the Dominated Convergence theorem yields $\xi = \mathrm{D}\mathcal{E}_{\mathrm{b}}(u)$, hence

$$\mathrm{D}\mathcal{E}_{\mathrm{b}}(u_\varepsilon) \rightharpoonup \mathrm{D}\mathcal{E}_{\mathrm{b}}(u) \quad \text{in } \mathrm{L}^2(\Omega_T). \tag{4.12}$$

Moreover, we have the following convergences for U_ε

$$\begin{aligned} U_\varepsilon &\overset{*}{\rightharpoonup} U && \text{in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Omega)),\\ \partial_\theta U_\varepsilon &\to 0 && \text{in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Sigma)), \end{aligned} \tag{4.13}$$

where the last convergence is due to $\beta > -1$ and $\varepsilon^{-(1+\beta)}\|\partial_\theta U_\varepsilon(t)\|^2_{\mathrm{L}^2(\Sigma)}$ being bounded. Depending on the regime for β we find a subsequence such that the tangential gradients of U_ε satisfy

$$\left.\begin{aligned} \nabla_\Gamma U_\varepsilon &\overset{*}{\rightharpoonup} \nabla_\Gamma U && \text{for } \beta = 1\\ \nabla_\Gamma U_\varepsilon &\to 0 && \text{for } \beta > 1 \end{aligned}\right\} \text{ in } \mathrm{L}^\infty(0,T;\mathrm{L}^2(\Sigma)). \tag{4.14}$$

Furthermore, we can assume that

$$\text{for } \alpha = 1: \qquad \dot{U}_\varepsilon \rightharpoonup \dot{U} \quad \text{and} \quad \mathrm{D}\overline{\mathcal{E}}_{\mathrm{s},\varepsilon}(U_\varepsilon) \rightharpoonup \Xi, \tag{4.15}$$

where we denoted by $\overline{\mathcal{E}}_{\mathrm{s},\varepsilon}$ the surface energy part such that $\overline{\mathcal{E}}_\varepsilon(u,U) = \mathcal{E}_{\mathrm{b}}(u) + \overline{\mathcal{E}}_{\mathrm{s},\varepsilon}(U)$. The limit $\Xi \in \mathrm{L}^2(\Sigma_T)$ is to be determined. For the remaining cases $\alpha < 1$ and $\alpha > 1$ we have

$$\left.\begin{aligned} \mathrm{D}\overline{\mathcal{E}}_{\mathrm{s},\varepsilon}(U_\varepsilon) &\to 0 && \text{for } \alpha < 1\\ \dot{U}_\varepsilon &\to 0 && \text{for } \alpha > 1 \end{aligned}\right\} \text{ in } \mathrm{L}^2(\Sigma_T). \tag{4.16}$$

Remark 4.1.4 *Let us remark that for all values of α and β considered here we can always find a (not relabeled) subsequence such that $\boldsymbol{u}_\varepsilon \to \boldsymbol{u}$* strongly *in $\mathrm{L}^2(0,T;\mathcal{H})$. Indeed, this follows from the fact that the traces $u_\varepsilon|_\Gamma = U_\varepsilon|_{\theta=0}$ converge strongly in $\mathrm{L}^2(0,T;\mathrm{L}^2(\Gamma))$ and $\partial_\theta U_\varepsilon \to 0$ strongly in $\mathrm{L}^2(\Sigma_T)$.*

Obviously, the energy functionals $\overline{\mathcal{E}}_\varepsilon$ blow up if the derivative with respect to θ does not vanish (for $\beta > 1$ the same holds for the tangential derivatives). Thus, we expect the limit problems to be defined on the subspace of functions that are constant in normal direction (and tangential direction for $\beta > 1$).

Let us consider the case $\beta \ge 1$ first: We define the reduced spaces $\mathcal{V}_{\mathrm{tang}}$, $\mathcal{V}_{\mathrm{const}}$ and their closures in $\mathcal{H}$ via

$$\mathcal{V}_{\mathrm{tang}} \stackrel{\mathrm{def}}{=} \{(u,U)\in\mathcal{V} \,\boldsymbol{:}\, \partial_\theta U = 0 \text{ a.e. in } \Sigma\}, \qquad \mathcal{H}_{\mathrm{tang}} \stackrel{\mathrm{def}}{=} \overline{\mathcal{V}_{\mathrm{tang}}}^{\mathcal{H}},$$
$$\mathcal{V}_{\mathrm{const}} \stackrel{\mathrm{def}}{=} \{(u,U)\in\mathcal{V} \,\boldsymbol{:}\, U = \text{const a.e. in } \Sigma\}, \qquad \mathcal{H}_{\mathrm{const}} \stackrel{\mathrm{def}}{=} \overline{\mathcal{V}_{\mathrm{const}}}^{\mathcal{H}}.$$

In the following theorem we prove the Mosco convergence of the energy functionals $\overline{\mathcal{E}}_\varepsilon$ in $\mathcal{V}$ in the regime $\beta \geq 1$.

Theorem 4.1.5 (Mosco convergence for $\beta \geq 1$) *For $\beta = 1$ the energy functionals $\overline{\mathcal{E}}_\varepsilon$ converge in the sense of Mosco to the limit functional $\overline{\mathcal{E}}_{\mathrm{tang}} : \mathcal{V} \to \mathbb{R}_\infty$ given by*

$$\overline{\mathcal{E}}_{\mathrm{tang}}(\boldsymbol{u}) = \begin{cases} \mathcal{E}_{\mathrm{b}}(u) + \int_\Sigma \left[\frac{A_{\mathrm{s}}}{2}|\nabla_\Gamma U|^2 + W_{\mathrm{s}}(U)\right]\mathrm{d}\mu & \text{if } \boldsymbol{u}\in\mathcal{V}_{\mathrm{tang}}, \\ +\infty & \text{otherwise.}\end{cases}$$

For $\beta > 1$ the Mosco limit of $\overline{\mathcal{E}}_\varepsilon$, denoted $\overline{\mathcal{E}}_{\mathrm{const}}$, is given by

$$\overline{\mathcal{E}}_{\mathrm{const}}(\boldsymbol{u}) = \begin{cases} \mathcal{E}_{\mathrm{b}}(u) + \int_\Sigma W_{\mathrm{s}}(U)\,\mathrm{d}\mu & \text{if } \boldsymbol{u}\in\mathcal{V}_{\mathrm{const}}, \\ +\infty & \text{otherwise.}\end{cases}$$

Proof: Here, we only consider the case $\beta = 1$. The result for the other case follows analogously.

Liminf estimate for weak convergence. For all sequences $\boldsymbol{u}_\varepsilon{=}(u_\varepsilon, U_\varepsilon) \rightharpoonup \boldsymbol{u}{=}(u,U)$ in $\mathcal{V}$ we have to show $\overline{\mathcal{E}}_{\mathrm{tang}}(\boldsymbol{u}) \leq \liminf_{\varepsilon\to0}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)$. Assuming that $\liminf_{\varepsilon\to0}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) < \infty$ due to the weak lower semicontinuity of the norm on $\mathcal{V}$ we necessarily have that $\boldsymbol{u}\in\mathcal{V}_{\mathrm{tang}}$.

The compact embedding $\mathcal{V} \subset\subset \mathrm{L}^q(\Omega)\times\mathrm{L}^q(\Sigma)$, where $q \in [1,\infty[$ for $d = 2$ and $q < 2d/(d{-}2)$ otherwise, yields the strong convergence $(u_\varepsilon, U_\varepsilon) \to (u,U)$ in $\mathrm{L}^q(\Omega)\times\mathrm{L}^q(\Sigma)$. Thus, using the growth conditions for W_{b} and W_{s} we conclude that

$$\int_\Omega W_{\mathrm{b}}(u_\varepsilon)\,\mathrm{d}x \to \int_\Omega W_{\mathrm{b}}(u)\,\mathrm{d}x \qquad \text{and} \qquad \int_\Sigma W_{\mathrm{s}}(U_\varepsilon)\,\mathrm{d}\mu \to \int_\Sigma W_{\mathrm{s}}(U)\,\mathrm{d}\mu.$$

As before, we denote the bulk and surface energy parts of $\overline{\mathcal{E}}_\varepsilon$ by $\mathcal{E}_{\mathrm{b}}$ and $\overline{\mathcal{E}}_{\mathrm{s},\varepsilon}$, such that $\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) = \mathcal{E}_{\mathrm{b}}(u_\varepsilon) + \overline{\mathcal{E}}_{\mathrm{s},\varepsilon}(U_\varepsilon)$. It holds that

$$\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \geq \mathcal{E}_{\mathrm{b}}(u_\varepsilon) + \int_\Sigma \left[\frac{A_{\mathrm{s}}}{2}\nabla_\Gamma U_\varepsilon\cdot\mathbb{B}_\varepsilon(y,\theta)\nabla_\Gamma U_\varepsilon + W_{\mathrm{s}}(U_\varepsilon)\right]\frac{\mathbb{J}_\varepsilon(y,\theta)}{\varepsilon}\,\mathrm{d}\mu.$$

Hence, by the uniform convergence of $\mathbb{B}_\varepsilon$ and $\mathbb{J}_\varepsilon/\varepsilon$ we obtain the lim inf estimate.

Limsup estimate for strongly converging recovery sequences. The construction of recovery sequences $\widehat{\boldsymbol{u}}_\varepsilon$ such that $\widehat{\boldsymbol{u}}_\varepsilon \to \boldsymbol{u}$ and $\overline{\mathcal{E}}_\varepsilon(\widehat{\boldsymbol{u}}_\varepsilon) \to \overline{\mathcal{E}}_{\mathrm{tang}}(\boldsymbol{u})$ is straightforward: For $\boldsymbol{u}\notin\mathcal{V}_{\mathrm{tang}}$ the result is trivial since $\overline{\mathcal{E}}_{\mathrm{tang}}(\boldsymbol{u}) = \infty$ and we may take $\widehat{\boldsymbol{u}}_\varepsilon = \boldsymbol{u}$ and argue as in the first step.

For $\boldsymbol{u}\in\mathcal{V}_{\mathrm{tang}}$ we can choose the constant sequence $\widehat{\boldsymbol{u}}_\varepsilon = \boldsymbol{u}$ since the derivative with respect to θ does not appear in $\overline{\mathcal{E}}_\varepsilon$, and we can conclude

$$\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}) = \mathcal{E}_{\mathrm{b}}(u) + \int_\Sigma \left[\frac{A_{\mathrm{s}}}{2}\nabla_\Gamma U \cdot \mathbb{B}_\varepsilon(y,\theta)\nabla_\Gamma U + W_{\mathrm{s}}(U)\right]\frac{\mathbb{J}_\varepsilon(y,\theta)}{\varepsilon}\,\mathrm{d}\mu \to \overline{\mathcal{E}}_{\mathrm{tang}}(\boldsymbol{u}),$$

where we used Lemma 4.1.2 again. □

The remaining case $\beta \in]-1,1[$ is more complicated since we lose the uniform coercivity of the energy functionals on $\mathcal{V}$. Hence, we have to work in the coarser topology of the bigger space $\mathcal{W}$ defined by

$$\mathcal{W} \stackrel{\mathrm{def}}{=} \left\{(u,U) \in \mathrm{H}^1(\Omega)\times\mathrm{L}^2(\Sigma)\ :\ \partial_\theta U \in \mathrm{L}^2(\Sigma),\ u|_\Gamma{=}U|_{\{\theta=0\}}\right\}.$$

Let us point out here that the existence of the derivative with respect to θ in $\mathrm{L}^2(\Sigma)$ suffices for the well-definedness of the trace on Γ since for arbitrary $U \in \mathrm{C}^\infty(\Sigma)$ it holds that

$$\|U|_{\{\theta=0\}}\|_{\mathrm{L}^2(\Gamma)} \le C\big(\|U\|_{\mathrm{L}^2(\Sigma)} + \|\partial_\theta U\|_{\mathrm{L}^2(\Sigma)}\big).$$

As before we introduce a reduced space of functions which are constant in normal direction

$$\mathcal{W}_{\mathrm{nodiff}} \stackrel{\mathrm{def}}{=} \{(u,U) \in \mathcal{W}\ :\ \partial_\theta U = 0 \text{ a.e. in } \Sigma\}.$$

Since the convergence of the surface variable U_ε is in general only weak in $\mathrm{L}^2(\Sigma)$ and the nonlinearity W_{s} is allowed to be nonconvex we have to replace W_{s} in the limit by its convex envelope, denoted W_{s}^{**} in the following (see e.g. [Bra02, Dal93]).

Theorem 4.1.6 (Γ-convergence for $-1 < \beta < 1$) *The functionals $\overline{\mathcal{E}}_\varepsilon$ Γ-converge on $\mathcal{W}$ to the limit functional $\overline{\mathcal{E}}_{\mathrm{nodiff}} : \mathcal{W} \to \mathbb{R}_\infty$ given by*

$$\overline{\mathcal{E}}_{\mathrm{nodiff}}(\boldsymbol{u}) = \begin{cases} \mathcal{E}_{\mathrm{b}}(u) + \int_\Sigma W_{\mathrm{s}}^{**}(U)\,\mathrm{d}\mu & \text{if } \boldsymbol{u} \in \mathcal{W}_{\mathrm{nodiff}}, \\ +\infty & \text{otherwise.}\end{cases}$$

Proof: *Liminf estimate for weak convergence.* Let $\boldsymbol{u}_\varepsilon{=}(u_\varepsilon,U_\varepsilon) \rightharpoonup \boldsymbol{u}{=}(u,U)$ in $\mathcal{W}$. By arguing as in Theorem 4.1.5 we can assume that $\boldsymbol{u} \in \mathcal{W}_{\mathrm{nodiff}}$ and $\sup_{0<\varepsilon<\varepsilon_0} \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) < \infty$. We have the estimate

$$\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \ge \mathcal{E}_{\mathrm{b}}(u_\varepsilon) + \int_\Sigma W_{\mathrm{s}}^{**}(U_\varepsilon)\frac{\mathbb{J}_\varepsilon(y,\theta)}{\varepsilon}\,\mathrm{d}\mu.$$

Applying $\liminf_{\varepsilon\to0}$ to both sides of the estimate and using the uniform convergence of $\mathbb{J}_\varepsilon/\varepsilon$ and the weak lower semicontinuity of $U \mapsto \int_\Sigma W^{**}(U)\,\mathrm{d}\mu$ on $\mathrm{L}^2(\Sigma)$ we conclude that $\liminf_{\varepsilon\to0} \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \ge \overline{\mathcal{E}}_{\mathrm{nodiff}}(\boldsymbol{u})$.

Limsup estimate for recovery sequences. Let $\boldsymbol{u} \in \mathcal{W}_{\mathrm{nodiff}}$ be such that $\overline{\mathcal{E}}_{\mathrm{nodiff}}(\boldsymbol{u}) < \infty$. By the density of $\mathcal{V}_{\mathrm{tang}}$ in $\mathcal{W}_{\mathrm{nodiff}}$ we can find a sequence $(\widehat{\boldsymbol{u}}_\varepsilon)_{\varepsilon>0} \subset \mathcal{V}_{\mathrm{tang}}$ such that $\widehat{\boldsymbol{u}}_\varepsilon \to \boldsymbol{u}$ (strongly) in $\mathcal{W}$ and $\varepsilon^\sigma\|\nabla_\Gamma\widehat{U}_\varepsilon\|^2_{\mathrm{L}^2(\Sigma)} \to 0$, where $\sigma = 1{-}\beta \in]0,2[$. Since $\widehat{\boldsymbol{u}}_\varepsilon = (\widehat{u}_\varepsilon, \widehat{U}_\varepsilon)$ converges strongly in $\mathcal{W}$ we can extract a (not relabeled) sequence such that

$\widehat{U}_\varepsilon(y,\theta) \to U(y,\theta)$ a.e. in Σ. Using Fatou's lemma we obtain

$$\begin{aligned}\limsup_{\varepsilon\to 0} \overline{\mathcal{E}}_\varepsilon(\widehat{\boldsymbol{u}}_\varepsilon) &\le \limsup_{\varepsilon\to 0} \left\{ \mathcal{E}_\mathrm{b}(\widehat{u}_\varepsilon) + \int_\Sigma \left[C\varepsilon^\sigma |\nabla_\Gamma \widehat{U}_\varepsilon|^2 + W_\mathrm{s}(U_\varepsilon) \right] \frac{\mathbb{J}_\varepsilon(y,\theta)}{\varepsilon} \,\mathrm{d}\mu \right\} \\ &\le \mathcal{E}_\mathrm{b}(u) + \int_\Sigma W_\mathrm{s}(U)\,\mathrm{d}\mu.\end{aligned}$$

The left-hand side, also known as Γ-limes superior (or upper Γ-limit), is weakly lower semicontinuous on $\mathcal{W}$ (see [Dal93, Bra02]). Hence, by taking the lower semicontinuous envelope on both sides we arrive at $\limsup_{\varepsilon\to 0} \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \le \overline{\mathcal{E}}_\mathrm{nodiff}(\boldsymbol{u})$. $\square$

Let us emphasize here that for W_s satisfying the growth condition (4.4) the directional derivative $\langle \mathrm{D}\overline{\mathcal{E}}_\mathrm{nodiff}(\boldsymbol{u}), \boldsymbol{w}\rangle$ is in general not well-defined for $\boldsymbol{u}, \boldsymbol{w} \in \mathcal{W}_\mathrm{nodiff}$ since we do not have the embedding $\mathcal{W}_\mathrm{nodiff} \subset \mathrm{L}^q(\Omega) \times \mathrm{L}^q(\Sigma)$ for $1 \le q < \infty$ for $d = 2$ and $1 \le q < 2d/(d-2)$ for $d \ge 3$. Thus, we restrict ourselves to the case of a quadratic potential, such that $W_\mathrm{s}(U) = \frac{\omega_\mathrm{s}}{2}|U|^2$ with $\omega_\mathrm{s} > 0$. In this much simpler case the (strongly converging) recovery sequences are given by $\widehat{\boldsymbol{u}}_\varepsilon$ in the proof of Theorem 4.1.6. Hence, $\overline{\mathcal{E}}_\varepsilon$ Mosco converges to $\overline{\mathcal{E}}_\mathrm{nodiff}$ in $\mathcal{W}$.

The limits for the dissipation potential $\overline{\Psi}_\varepsilon$ and the dual dissipation potentials $\overline{\Psi}^*_\varepsilon$ for the cases $\alpha{=}1$, $\alpha{>}1$ and $\alpha{<}1$ are easily computed. Note that for the last two cases the uniform coercivity of $\overline{\Psi}^*_\varepsilon$ and $\overline{\Psi}_\varepsilon$ on $\mathcal{H}^*$ and $\mathcal{H}$, respectively, is lost.

For the nondegenerate case $\alpha{=}1$ we have the convergence

$$\overline{\Psi}_\varepsilon \xrightarrow{\mathrm{M}} \overline{\Psi}_\mathrm{dyn} \qquad \text{with} \qquad \overline{\Psi}_\mathrm{dyn}(\dot{\boldsymbol{u}}) = \int_\Omega \frac{\tau_\mathrm{b}}{2}|\dot{u}|^2\,\mathrm{d}x + \int_\Sigma \frac{\tau_\mathrm{s}}{2}|\dot{U}|^2\,\mathrm{d}\mu$$

while for the other two cases (the slow and the fast evolution cases, see discussion in Section 4.1.3) it holds

$$\begin{aligned}\alpha > 1:\ & \overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}) \to \overline{\Psi}_\mathrm{slow}(\dot{\boldsymbol{u}}) \quad \text{with} \quad \overline{\Psi}_\mathrm{slow}(\dot{u},\dot{U}) = \begin{cases} \int_\Omega \frac{\tau_\mathrm{b}}{2}|\dot{u}|^2\,\mathrm{d}x & \text{if } \dot{U}{=}0, \\ \infty & \text{else,} \end{cases} \\ \alpha < 1:\ & \overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}) \to \overline{\Psi}_\mathrm{fast}(\dot{\boldsymbol{u}}) \quad \text{with} \quad \overline{\Psi}_\mathrm{fast}(\dot{u},\dot{U}) = \int_\Omega \frac{\tau_\mathrm{b}}{2}|\dot{u}|^2\,\mathrm{d}x.\end{aligned}$$

The Legendre transforms are easily computed as

$$\overline{\Psi}^*_\mathrm{slow}(\xi,\Xi) = \int_\Omega \frac{\tau_\mathrm{b}^{-1}}{2}|\xi|^2\,\mathrm{d}x \quad \text{and} \quad \overline{\Psi}^*_\mathrm{fast}(\xi,\Xi) = \begin{cases} \int_\Omega \frac{\tau_\mathrm{b}^{-1}}{2}|\xi|^2\,\mathrm{d}x & \text{if } \Xi = 0, \\ \infty & \text{else.} \end{cases}$$

We see that the limits for $\overline{\Psi}_\varepsilon$ correspond to the observations made in Section 4.1.3. For $\alpha{>}1$ we obtain the static condition $\dot{U} = 0$, i.e., fixed (boundary) evolution. While for $\alpha{<}1$ the condition $\Xi = 0$ for the thermodynamically conjugated driving force means that the (boundary-)system is in equilibrium.

Passing to the limit in the energy balance

In this subsection we focus on the energy-dissipation formulation (4.9) and show that the limit $\boldsymbol{u} = (u, U)$ in (4.11)–(4.15) is a solution of the limit system $(\overline{\mathcal{E}}_0, \overline{\Psi}_0)$ with $\overline{\mathcal{E}}_0 = \overline{\mathcal{E}}_{\text{tang}}, \overline{\mathcal{E}}_{\text{const}}, \overline{\mathcal{E}}_{\text{nodiff}}$ and $\overline{\Psi}_0 = \overline{\Psi}_{\text{slow}}, \overline{\Psi}_{\text{dyn}}$. In particular, we do not treat the case $\overline{\Psi}_0 = \overline{\Psi}_{\text{fast}}$ since in this limit case the chain rule is not available. Hence, the abstract framework discussed in Chapter 2 does not apply in this case, and we are not able to characterize the limit $\boldsymbol{u}$ as a solution of a corresponding force balance formulation. However, we show in the following subsection that for λ-convex energies the (EVI$_\lambda$)-formulation can be used instead.

In particular, we show in this subsection that

$$\liminf_{\varepsilon\to 0} \left\{ \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(t)) + \int_0^t \left[\overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}_\varepsilon) + \overline{\Psi}_\varepsilon^*(-\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon))\right] \mathrm{d}s \right\} \\ \geq \overline{\mathcal{E}}_0(\boldsymbol{u}(t)) + \int_0^t \left[\overline{\Psi}_0(\dot{\boldsymbol{u}}) + \overline{\Psi}_0^*(-\mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}))\right] \mathrm{d}s.$$

Here, and in the following, to cover all limit cases we use the the notation $\mathcal{V}_0 = \mathcal{V}_{\text{tang}}, \mathcal{V}_{\text{const}}$ and $\mathcal{W}_{\text{nodiff}}$ when we refer to the domains of the corresponding limit energy functionals $\overline{\mathcal{E}}_0 = \overline{\mathcal{E}}_{\text{tang}}$, etc.

Remark 4.1.7 *In order to pass to the limit we use the pointwise (in time) weak convergence of the solutions in the space $\mathcal{V}$ (resp. $\mathcal{W}$), i.e., $\boldsymbol{u}_\varepsilon(t) \rightharpoonup \boldsymbol{u}(t)$ in $\mathcal{V}$ (resp. $\mathcal{W}$).*

Indeed, let $\mathcal{V}_{\text{weak}}$ denote the space $\mathcal{V}$ endowed with the weak topology. Then, from the continuous embedding $\mathrm{L}^\infty(0,T;\mathcal{V}) \cap \mathrm{H}^1(0,T;\mathcal{H}) \subset \mathrm{C}([0,T];\mathcal{V}_{\text{weak}})$ *(see e.g. [Rou05, Sect. 8.3]) we obtain $\boldsymbol{u}_\varepsilon(t) \rightharpoonup \boldsymbol{u}(t)$ in $\mathcal{V}$ (the same holds for $\mathcal{V}$ replaced by $\mathcal{W}$). This can be seen by means of a simple contradiction argument.*

Following the ideas in [SaS04] we define for $\boldsymbol{u}_\varepsilon$ the *energy excess* $\overline{\mathcal{D}} : [0,T] \to [0,\infty]$ by

$$\overline{\mathcal{D}}_\varepsilon(t) = \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(t)) - \overline{\mathcal{E}}_0(\boldsymbol{u}(t)), \qquad \overline{\mathcal{D}}(t) = \limsup_{\varepsilon\to 0} \overline{\mathcal{D}}_\varepsilon(t) \geq 0.$$

We call $\boldsymbol{u}_\varepsilon$ *well-prepared initially* if $\overline{\mathcal{D}}(0) = 0$. Notably, this can be translated into asking that $\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(0)) \to \overline{\mathcal{E}}_0(\boldsymbol{u}(0))$, i.e., the initial energies converge.

The additional conditions for the convergence of the gradient flow given in [SaS04] can be directly translated in our case to

1. *(Lower Bound)* There exists $f \in \mathrm{L}^1(0,T)$ such that for every $t \in [0,T]$
$$\liminf_{\varepsilon\to 0} \int_0^t \overline{\Psi}_\varepsilon(\dot{\boldsymbol{u}}_\varepsilon)\,\mathrm{d}s \geq \int_0^t \left[\overline{\Psi}_0(\dot{\boldsymbol{u}}) - f(s)\overline{\mathcal{D}}(s)\right] \mathrm{d}s. \tag{4.17}$$

2. *(Construction)* There exists a locally bounded function g on $[0,T]$ such that for any $t_0 \in \,]0,T[$ and any smooth curve $\widehat{\boldsymbol{u}} : \,]t_0{-}\rho, t_0{+}\rho[\, \to \mathcal{V}_0$ satisfying $\widehat{\boldsymbol{u}}(t_0){=}\boldsymbol{u}(t_0)$ there

exists a $\widehat{\boldsymbol{u}}_\varepsilon :]t_0{-}\rho, t_0{+}\rho[\to \mathcal{V}$ such that $\widehat{\boldsymbol{u}}_\varepsilon(t_0){=}\boldsymbol{u}_\varepsilon(t_0)$ and

$$\limsup_{\varepsilon\to 0} \overline{\Psi}_\varepsilon(\dot{\widehat{\boldsymbol{u}}}_\varepsilon(t_0)) \le \overline{\Psi}_0(\dot{\widehat{\boldsymbol{u}}}(t_0)) + g(t_0)\overline{\mathcal{D}}(t_0), \tag{4.18a}$$

$$\liminf_{\varepsilon\to 0} -\frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathcal{E}}_\varepsilon(\widehat{\boldsymbol{u}}_\varepsilon)|_{t=t_0} \ge -\frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathcal{E}}_0(\widehat{\boldsymbol{u}})|_{t=t_0} - g(t_0)\overline{\mathcal{D}}(t_0). \tag{4.18b}$$

The energy excess $\overline{\mathcal{D}}$ should be interpreted as a small perturbation. It is shown in [SaS04] that $\overline{\mathcal{D}} \equiv 0$ holds using Gronwall's lemma. However, in the present case we can simply take $f = g = 0$.

While the first condition in (4.17) asks for a liminf estimate for the (integrated) dissipation potential $\overline{\Psi}_\varepsilon$, the second condition in (4.18) can be interpreted as a liminf estimate for the dual dissipation potential along the derivative of the energy functionals. Indeed, summing up (4.18a) and (4.18b) we arrive at

$$\begin{aligned}\liminf_{\varepsilon\to 0} \overline{\Psi}^*_\varepsilon(-\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)) &\ge \liminf_{\varepsilon\to 0}\Big[-\langle \mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon), \dot{\widehat{\boldsymbol{u}}}_\varepsilon\rangle - \overline{\Psi}_\varepsilon(\dot{\widehat{\boldsymbol{u}}}_\varepsilon)\Big] \\ &\ge -\langle \mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}), \dot{\widehat{\boldsymbol{u}}}\rangle - \overline{\Psi}_0(\dot{\widehat{\boldsymbol{u}}}).\end{aligned}$$

Taking the supremum with respect to $\dot{\widehat{\boldsymbol{u}}}$ yields the dissipation potential $\overline{\Psi}^*_{\mathrm{dyn}}(-\mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}))$ in the right-hand side.

Let us point out that the limit system considered in [SaS04] is finite dimensional. Therefore, we have to adapt the results for our purpose. In particular, we have to show that the Gâteaux derivative of the limit energy functional is well-defined in $\mathcal{H}$.

The main result for $\overline{\mathcal{E}}_0 = \overline{\mathcal{E}}_{\mathrm{tang}}, \overline{\mathcal{E}}_{\mathrm{const}}$ and $\overline{\mathcal{E}}_{\mathrm{nodiff}}$ and $\overline{\Psi}_0 = \overline{\Psi}_{\mathrm{dyn}}$ reads as follows:

Theorem 4.1.8 (Convergence for $\beta > -1$ and $\alpha{=}1$) *Let $\boldsymbol{u}_\varepsilon$ be a family of solutions of* (4.9) *converging as in* (4.11)–(4.15) *to a limit $\boldsymbol{u}$. If $\overline{\mathcal{D}}(0) = 0$, then $\boldsymbol{u}$ is the solution of the gradient flow for $\overline{\mathcal{E}}_0$ and $\overline{\Psi}_{\mathrm{dyn}}$, i.e.,*

$$\overline{\mathcal{E}}_0(\boldsymbol{u}(t)) + \int_0^t \Big[\overline{\Psi}_{\mathrm{dyn}}(\dot{\boldsymbol{u}}) + \overline{\Psi}^*_{\mathrm{dyn}}(-\mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}))\Big]\,\mathrm{d}s \le \overline{\mathcal{E}}_0(\boldsymbol{u}(0)). \tag{4.19}$$

Proof: First, we note that the weak convergence $\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon) \rightharpoonup \boldsymbol{\xi} = (\mathrm{D}\mathcal{E}_{\mathrm{b}}(u), \Xi)$ in $\mathrm{L}^2(0,T;\mathcal{H}^*)$ (see (4.12) and (4.15)) implies $\mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}) \in \mathrm{L}^2(0,T;\mathcal{H}_0^*)$, where $\mathcal{H}_0 = \overline{\mathcal{V}_0}^{\mathcal{H}}$. Indeed, for an arbitrary $\widehat{\boldsymbol{v}} \in \mathrm{L}^2(0,T;\mathcal{V}_0\cap\mathcal{V})$ we get the convergence

$$\int_0^T \langle \mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon), \widehat{\boldsymbol{v}}\rangle\,\mathrm{d}t \to \int_0^T \langle \mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}), \widehat{\boldsymbol{v}}\rangle\,\mathrm{d}t = \int_0^T \langle \boldsymbol{\xi}, \widehat{\boldsymbol{v}}\rangle\,\mathrm{d}t.$$

Here, we used the continuity properties of the associated Nemytskii operators $u \mapsto W_{\mathrm{b}}'(u)$ and $U \mapsto W_{\mathrm{s}}'(U)$, respectively (see e.g. [Rou05, Sect. 8.6]).

Moreover, we see that an arbitrary $\widehat{\boldsymbol{v}} \in \mathrm{L}^2(0,T;\mathcal{V}_0\cap\mathcal{V})$ satisfies the conditions (4.18a) and (4.18b): We easily check that $\int_0^t \overline{\Psi}_\varepsilon(\widehat{\boldsymbol{v}})\,\mathrm{d}s \to \int_0^t \overline{\Psi}_{\mathrm{dyn}}(\widehat{\boldsymbol{v}})\,\mathrm{d}s$ holds and conclude that

$$\liminf_{\varepsilon\to 0}\int_0^t \overline{\Psi}^*_\varepsilon(-\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon))\,\mathrm{d}s \geq \liminf_{\varepsilon\to 0}\int_0^t \Big[-\langle \mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon),\widehat{\boldsymbol{v}}\rangle - \overline{\Psi}_\varepsilon(\widehat{\boldsymbol{v}})\Big]\,\mathrm{d}s$$
$$= \int_0^t \Big[-\langle \mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}),\widehat{\boldsymbol{v}}\rangle - \overline{\Psi}_{\mathrm{dyn}}(\widehat{\boldsymbol{v}})\Big]\,\mathrm{d}s.$$

Taking the supremum over all $\widehat{\boldsymbol{v}} \in \mathrm{L}^2(0,T;\mathcal{H}_0)$ we arrive at the liminf estimate for the dual dissipation potential.

The Mosco convergence of the energy functionals and Remark 4.1.7 lead together with the liminf estimate for $\overline{\Psi}_\varepsilon$ to the lower energy estimate

$$\overline{\mathcal{E}}_0(\boldsymbol{u}(t)) + \int_0^t \big[\overline{\Psi}_{\mathrm{dyn}}(\dot{\boldsymbol{u}}) + \overline{\Psi}^*_{\mathrm{dyn}}(-\mathrm{D}\overline{\mathcal{E}}_0(\boldsymbol{u}))\big]\,\mathrm{d}s \leq \overline{\mathcal{E}}_0(\boldsymbol{u}(0)),$$

which is actually an equality due to the chain rule for $t \mapsto \overline{\mathcal{E}}_0(\boldsymbol{u}(t))$ and the characterization of the Legendre transform. □

The derivation of the corresponding energy balance for $\overline{\Psi}_0 = \overline{\Psi}_{\mathrm{slow}}$ is remarkably easier and follows by the same arguments as in the proof of Theorem 4.1.8 with $\widehat{\boldsymbol{v}} = 0$.

Theorem 4.1.9 (Convergence for $\beta{>}{-1}$ and $\alpha > 1$) *Let $\boldsymbol{u}_\varepsilon$ be a family of solutions of* (4.9) *converging as in* (4.11)–(4.15) *to a limit $\boldsymbol{u}$. If $\overline{\mathcal{D}}(0) = 0$ then $\boldsymbol{u}$ is the solution of the gradient flow for $\overline{\mathcal{E}}_0$ and $\overline{\Psi}_{\mathrm{slow}}$, i.e.,*

$$\mathcal{E}_{\mathrm{b}}(u(t)) + \int_0^t \big[\overline{\Psi}_{\mathrm{b}}(\dot{u}) + \overline{\Psi}^*_{\mathrm{b}}(-\mathrm{D}\mathcal{E}_{\mathrm{b}}(u))\big]\,\mathrm{d}s = \mathcal{E}_{\mathrm{b}}(u(0)),$$

where $\mathcal{E}_{\mathrm{b}}$ and $\overline{\Psi}_{\mathrm{b}}$ denote the bulk part of the limit energy and dissipation potential, such that $\overline{\mathcal{E}}_0(\boldsymbol{u}) = \mathcal{E}_{\mathrm{b}}(u) + \overline{\mathcal{E}}_{\mathrm{s},0}(U)$ and $\overline{\Psi}_{\mathrm{slow}}(\dot{\boldsymbol{u}}) = \overline{\Psi}_{\mathrm{b}}(\dot{u})$.

Passing to the limit in the evolutionary variational inequality

In order to treat the degenerate case $\alpha < 1$ we assume additionally that the potentials $W_{\mathrm{b/s}}$ are λ-convex on $\mathbb{R}$, i.e., $u \mapsto W_{\mathrm{b/s}}(u) - \frac{\lambda}{2}|u|^2$ is convex. This means in particular, that $W = W_{\mathrm{b/s}}$ satisfies the convexity estimate

$$\forall\, u,\widetilde{u} \in \mathbb{R}: \qquad W(\widetilde{u}) \geq W(u) + W'(u)\cdot(\widetilde{u}-u) + \frac{\lambda}{2}|\widetilde{u}-u|^2.$$

Note that λ does not have to be positive and therefore W_{b} and W_{s} are in general nonconvex, e.g. the double well potential $u \mapsto \frac{1}{4}(1-u^2)^2$ is λ-convex with $\lambda = -1$. Moreover, every $W \in \mathrm{C}^{1,1}(\mathbb{R})$ is λ-convex with $\lambda = -\mathrm{Lip}(W')$.

As a consequence, the energy functionals $\overline{\mathcal{E}}_\varepsilon$ satisfy for all $\widetilde{\boldsymbol{u}} = (\widetilde{u},\widetilde{U}), \boldsymbol{u} = (u,U) \in \mathcal{V}$ the estimate

$$\overline{\mathcal{E}}_\varepsilon(\widetilde{\boldsymbol{u}}) \geq \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}) + \langle \mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}),\widetilde{\boldsymbol{u}}-\boldsymbol{u}\rangle + \frac{\lambda}{2}\big\{\|u-\widetilde{u}\|^2_{\mathrm{L}^2(\Omega)} + \textstyle\int_\Sigma |U-\widetilde{U}|^2 \frac{\mathbb{J}_\varepsilon}{\varepsilon}\,\mathrm{d}\mu\big\}. \tag{4.20}$$

In view of Chapter 3 we introduce the distance $\mathbf{d}_\varepsilon : \mathcal{H} \times \mathcal{H} \to [0, \infty[$ induced by the metric tensor $\overline{\mathcal{G}}_\varepsilon = \mathrm{D}\overline{\Psi}_\varepsilon : \mathcal{H} \to \mathcal{H}^*$ as

$$\mathbf{d}_\varepsilon(\widetilde{\boldsymbol{u}}, \boldsymbol{u})^2 = \langle \overline{\mathcal{G}}_\varepsilon(\widetilde{\boldsymbol{u}}-\boldsymbol{u}), \widetilde{\boldsymbol{u}}-\boldsymbol{u}\rangle = \tau_\mathrm{b}\|u-\widetilde{u}\|^2_{\mathrm{L}^2(\Omega)} + \int_\Sigma \tau_\mathrm{s}\varepsilon^{-\alpha}\big|U-\widetilde{U}\big|^2 \mathbb{J}_\varepsilon \,\mathrm{d}\mu.$$

Hence, by defining $\lambda_\varepsilon \stackrel{\mathrm{def}}{=} \min\{\frac{\lambda}{\tau_\mathrm{b}}, \frac{\lambda\varepsilon^{\alpha-1}}{\tau_\mathrm{s}}\}$ we obtain from (4.20) that $\overline{\mathcal{E}}_\varepsilon$ is geodesically λ_ε-convex with respect to $\mathbf{d}_\varepsilon$. However, in the case $\alpha < 1$ we have $\varepsilon^{\alpha-1} \to \infty$. Hence, if $\lambda < 0$ we unfortunately have $\lambda_\varepsilon \to -\infty$ for $\varepsilon \to 0$. To circumvent this problem we directly use the gradient flow equation $\overline{\mathcal{G}}_\varepsilon \dot{\boldsymbol{u}}_\varepsilon = -\mathrm{D}\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon)$ in (4.20) to obtain the evolutionary variational inequality

$$\forall\, t \in [0,T],\, \widetilde{\boldsymbol{u}} \in \mathcal{V}: \quad \langle \overline{\mathcal{G}}_\varepsilon \dot{\boldsymbol{u}}_\varepsilon(t), \boldsymbol{u}_\varepsilon(t)-\widetilde{\boldsymbol{u}}\rangle + \overline{\Lambda}_\varepsilon(\boldsymbol{u}(t)-\widetilde{\boldsymbol{u}}) + \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(t)) \leq \overline{\mathcal{E}}(\widetilde{\boldsymbol{u}}), \qquad (4.21)$$

where $\overline{\Lambda}(\boldsymbol{v}) = \lambda\|v\|^2_{\mathrm{L}^2(\Omega)} + \lambda \int_\Sigma |V|^2 \frac{\mathbb{J}_\varepsilon}{\varepsilon}\,\mathrm{d}\mu$.

Let us remark here that this can be seen as using the separate geodesic λ-convexity of the bulk and surface energy $\mathcal{E}_\mathrm{b}$ and $\overline{\mathcal{E}}_{\mathrm{s},\varepsilon}$ with respect to the norm on $\mathrm{L}^2(\Omega)$ and $\mathrm{L}^2(\Sigma)$, respectively.

Lemma 4.1.3 shows that in the regime $\alpha < 1$ we do not have estimates for the time derivative $\dot{\boldsymbol{u}}_\varepsilon$ and therefore are not able to make statements about the pointwise in time convergence of $\boldsymbol{u}_\varepsilon$. Therefore, we integrate (4.21) with respect to $t \in [0,T]$. Though, the integrated inequality is in general difficult to treat since the Γ-convergence of the time-integrated functionals $\boldsymbol{u} \mapsto \int_0^T \overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}(t))\,\mathrm{d}t$ is (in general) not trivial. We refer to [Ste08a, Sal84] for the following result.

Proposition 4.1.10 *Let $\mathcal{F}_\varepsilon$ denote a sequence of weakly lower semicontinuous functionals on a reflexive and separable Banach space $\mathcal{X}$ satisfying the liminf estimate for the weak convergence in $\mathcal{X}$. Moreover, let $w_\varepsilon \rightharpoonup w$ (weakly-* if $p = \infty$) in $\mathrm{L}^p(0,T;\mathcal{X})$, then*

$$\int_0^T \mathcal{F}_0(w(t))\,\mathrm{d}t \leq \liminf_{\varepsilon\to 0} \int_0^T \mathcal{F}_\varepsilon(w_\varepsilon(t))\,\mathrm{d}t.$$

The main result for the case $\overline{\Psi}_0 = \overline{\Psi}_\mathrm{fast}$ reads as follows.

Theorem 4.1.11 (Convergence for $\beta{>}{-1}$ and $\alpha{<}1$) *Let $\boldsymbol{u}_\varepsilon$ be a family of solutions of the evolutionary variational inequality* (4.21) *converging as in* (4.11)–(4.15) *to the limit $\boldsymbol{u}$. Then, $\boldsymbol{u}$ is the solution of the following evolutionary variational inequality for $\overline{\mathcal{E}}_0$ and $\overline{\Psi}_\mathrm{fast}$*

$$\int_0^T \Big[\tau_\mathrm{b}\langle \dot{u}, u-\widetilde{u}\rangle + \overline{\mathcal{E}}_0(\boldsymbol{u}) + \overline{\Lambda}_0(\boldsymbol{u}-\widetilde{\boldsymbol{u}})\Big]\,\mathrm{d}t \leq \int_0^T \overline{\mathcal{E}}_0(\widetilde{\boldsymbol{u}})\,\mathrm{d}t \qquad (4.22)$$

for all $\widetilde{\boldsymbol{u}} \in \mathrm{L}^2(0,T;\mathcal{V}_0)$, where $\overline{\Lambda}_0(\boldsymbol{u}) = \int_\Omega \frac{\lambda_\mathrm{b}}{2}|u|^2\,\mathrm{d}x + \int_\Sigma \frac{\lambda_\mathrm{s}}{2}|U|^2\,\mathrm{d}\mu$.

Proof: For each of the different limit cases of the energy functionals $\overline{\mathcal{E}}_\varepsilon$ let us choose a $\widetilde{\boldsymbol{u}}{=}(\widetilde{u}, \widetilde{U}) \in \mathrm{L}^2(0,T;\mathcal{V})$ such that $\widetilde{\boldsymbol{u}}(t) \in \mathrm{dom}(\overline{\mathcal{E}}_0)$. Arguing as in the proof of Theorems

4.1.5 and 4.1.6 and using Proposition 4.1.10 we obtain

$$\liminf_{\varepsilon\to 0}\left\{\int_0^T \left[\overline{\mathcal{E}}_\varepsilon(\boldsymbol{u}_\varepsilon(t)) - \overline{\mathcal{E}}_\varepsilon(\widetilde{\boldsymbol{u}}(t))\right]\mathrm{d}t\right\} \geq \int_0^T \left[\overline{\mathcal{E}}_0(\boldsymbol{u}(t)) - \overline{\mathcal{E}}_0(\widetilde{\boldsymbol{u}}(t))\right]\mathrm{d}t.$$

Moreover, from the estimates in Lemma 4.1.3 we infer that $\dot{u}_\varepsilon \rightharpoonup \dot{u}$ in $\mathrm{L}^2(\Omega_T)$ and $\varepsilon^{1-\alpha}\dot{U}_\varepsilon \to 0$ in $\mathrm{L}^2(\Sigma_T)$. Hence, we have that

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_0^T \langle \overline{\mathcal{G}}_\varepsilon \dot{\boldsymbol{u}}_\varepsilon, \boldsymbol{u}_\varepsilon - \widetilde{\boldsymbol{u}}\rangle\,\mathrm{d}t &= \lim_{\varepsilon\to 0}\int_0^T \tau_{\mathrm{b}}\langle \dot{u}_\varepsilon, u_\varepsilon - \widetilde{u}\rangle + \tau_{\mathrm{s}}\varepsilon^{1-\alpha}\Big\langle \dot{U}_\varepsilon, (U_\varepsilon - \widetilde{U})\frac{\mathbb{J}_\varepsilon}{\varepsilon}\Big\rangle\,\mathrm{d}t\\ &= \int_0^T \tau_{\mathrm{b}}\langle \dot{u}, u - \widetilde{u}\rangle\,\mathrm{d}t.\end{aligned}$$

Thus, using also that $\lim_{\varepsilon\to 0}\int_0^T \overline{\Lambda}_\varepsilon(\boldsymbol{u}_\varepsilon - \widetilde{\boldsymbol{u}})\,\mathrm{d}t = \int_0^T \overline{\Lambda}_0(\boldsymbol{u} - \widetilde{\boldsymbol{u}})\,\mathrm{d}t$ we finally obtain (4.22). $\square$

4.1.4 Discussion of the limit models

In this section we show that the limit models obtained in Section 4.1.3 can be reduced to a real bulk–surface evolutionary system in $\overline{\Omega}$. The main observation is that for a pair (u, U) in $\mathcal{V}_{\mathrm{tang}}, \mathcal{V}_{\mathrm{const}}$ or $\mathcal{W}_{\mathrm{nodiff}}$ we can characterize U by a function defined only on the boundary $\Gamma = \partial\Omega$. More precisely, these spaces are isomorph to the spaces V_{tang}, V_{const} and W_{nodiff} given by

$$\begin{aligned}V_{\mathrm{tang}} &\stackrel{\mathrm{def}}{=} \Big\{(u, U) \in \mathrm{H}^1(\Omega)\times\mathrm{H}^1(\Gamma) \,:\, u|_\Gamma = U\Big\},\\ V_{\mathrm{const}} &\stackrel{\mathrm{def}}{=} \Big\{(u, U) \in \mathrm{H}^1(\Omega)\times\mathbb{R}^{N_\Gamma} \,:\, u|_{\Gamma_i} = U^i,\ i = 1, \dots, N_\Gamma\Big\},\\ W_{\mathrm{nodiff}} &\stackrel{\mathrm{def}}{=} \Big\{(u, U) \in \mathrm{H}^1(\Omega)\times\mathrm{L}^2(\Gamma) \,:\, u|_\Gamma = U\Big\}\end{aligned}$$

where $N_\Gamma \in \mathbb{N}$ is the number of connected components $\Gamma_i \subset \Gamma$. We denote by H_{tang}, H_{const} and H_{nodiff} the closures of the spaces above with respect to the L^2-norm, such that

$$H_{\mathrm{tang}} = H_{\mathrm{nodiff}} = \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Gamma) \qquad \text{and} \qquad H_{\mathrm{const}} = \mathrm{L}^2(\Omega) \times \mathbb{R}^{N_\Gamma}.$$

With these characterizations the energy functionals $\overline{\mathcal{E}}_{\mathrm{tang}}$ and $\overline{\mathcal{E}}_{\mathrm{nodiff}}$ can be reduced by integration over the variable $\theta \in]0, 1[$ while for $\overline{\mathcal{E}}_{\mathrm{const}}$ we integrate over y as well. The reduced energy functionals, denoted $\mathcal{E}_{\mathrm{tang}}, \mathcal{E}_{\mathrm{const}}$ and $\mathcal{E}_{\mathrm{nodiff}}$ are then given by

$$\begin{aligned}\mathcal{E}_{\mathrm{tang}}(u, U) &\stackrel{\mathrm{def}}{=} \mathcal{E}_{\mathrm{b}}(u) + \int_\Gamma \left[\frac{A_{\mathrm{s}}}{2}|\nabla_\Gamma U|^2 + W_{\mathrm{s}}(U)\right]\mathrm{d}\Gamma,\\ \mathcal{E}_{\mathrm{const}}(u, U) &\stackrel{\mathrm{def}}{=} \mathcal{E}_{\mathrm{b}}(u) + |\Gamma|\sum_{i=1}^{N_\Gamma} W_{\mathrm{s}}(U^i),\\ \mathcal{E}_{\mathrm{nodiff}}(u, U) &\stackrel{\mathrm{def}}{=} \mathcal{E}_{\mathrm{b}}(u) + \frac{\omega_{\mathrm{s}}}{2}\|U\|^2_{\mathrm{L}^2(\Gamma)},\end{aligned}$$

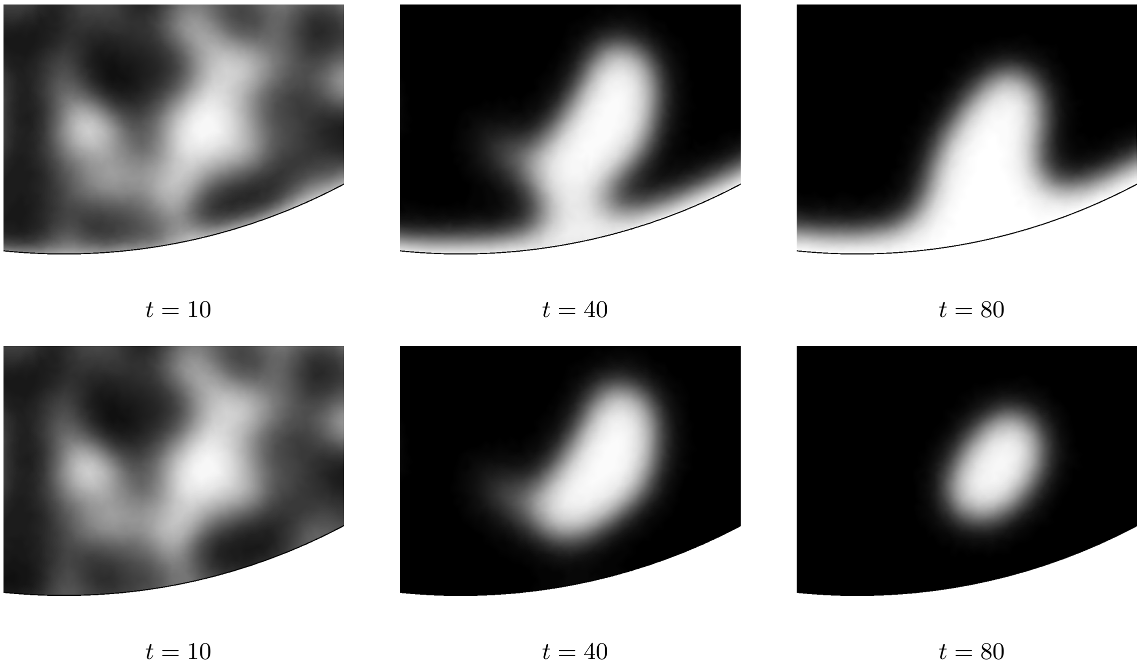

Figure 4.2: Detail of two solutions of equation (4.23) near the boundary with dynamic boundary condition (4.24) (**top**) and Neumann boundary condition (**bottom**) for subsequent times.

where in each case $\mathcal{E}_\mathrm{b}(u) = \int_\Omega [\frac{A_\mathrm{b}}{2}|\nabla u|^2 + W_\mathrm{b}(u)]\,\mathrm{d}x$ denotes the bulk energy. Starting with the case $\alpha = 1$ we see that the limit energy balance in (4.19) can be written in terms of $\mathcal{E}_0 \in \{\mathcal{E}_\mathrm{tang}, \mathcal{E}_\mathrm{const}, \mathcal{E}_\mathrm{nodiff}\}$ and the dissipation potential Ψ_dyn. Here, in slight abuse of notation, Ψ_dyn is for each of the energy functionals $\mathcal{E}_\mathrm{tang}$, $\mathcal{E}_\mathrm{const}$ and $\mathcal{E}_\mathrm{nodiff}$ defined on the spaces H_tang, H_const and H_nodiff and obtained as before via integration with respect to the variable θ or (y, θ), respectively. Thus, the reduced energy balance reads

$$\mathcal{E}_0(u(t), U(t)) + \int_0^t \left[\Psi_\mathrm{dyn}(\dot{u}, \dot{U}) + \Psi^*_\mathrm{dyn}(-\mathrm{D}\mathcal{E}_0(u, U))\right] \mathrm{d}s = \mathcal{E}_0(u(0), U(0)).$$

To highlight the structure of the limit systems we now write down the corresponding force balance equation written in terms of the reduced energy and dissipation functional. It consists of two equations for the bulk and the surface variable u and $U = u|_\Gamma$, respectively. Using the chain rule and the Fenchel equivalences we obtain

$$\begin{pmatrix} \tau_\mathrm{b}\dot{u} + \mathrm{D}_u\mathcal{E}_0(u, U) \\ \tau_\mathrm{s}\dot{U} + \mathrm{D}_U\mathcal{E}_0(u, U) \end{pmatrix} = 0.$$

For each of the energy functionals the first equation is formally equivalent to the well-known Allen–Cahn equation in $[0,T]\times\Omega$

$$\tau_{\mathrm{b}}\dot{u} - A_{\mathrm{b}}\Delta u + W_{\mathrm{b}}'(u) = 0. \tag{4.23}$$

This equation is coupled to the boundary evolution of $u|_\Gamma = U$, which for the energy functional $\mathcal{E}_{\mathrm{tang}}$ (limit case for $\beta = 1$) is described by

$$\tau_{\mathrm{s}}\dot{U} - A_{\mathrm{s}}\Delta_\Gamma U + A_{\mathrm{b}}\tfrac{\partial u}{\partial\nu} + W_{\mathrm{s}}'(U) = 0. \tag{4.24}$$

Hence, we obtain the surface Allen-Cahn equation with a contribution given by the conormal derivative of the bulk variable u. The system (4.23) & (4.24) was studied in [SpW10] regarding existence and uniqueness of global solutions, as well as asymptotic behavior and the existence of a global attractor.

In the limit case for $\beta > 1$ the limit energy functional is given by $\mathcal{E}_{\mathrm{const}}$, and we obtain a simpler boundary condition, which consists of a system of ordinary differential equations for each of the connected components Γ_i of the boundary Γ, namely

$$\tau_{\mathrm{s}}\dot{U}^i + A_{\mathrm{b}}\big[\tfrac{\partial u}{\partial\nu}\big]_i + W_{\mathrm{s}}'(U^i) = 0, \tag{4.25}$$

where $[g]_i \overset{\mathrm{def}}{=} \frac{1}{|\Gamma_i|}\int_{\Gamma_i} g\,\mathrm{d}\Gamma$ denotes the mean value of $g:\Gamma_i\to\mathbb{R}$ over $\Gamma_i\subset\Gamma$.

Finally, for $\mathcal{E}_0 = \mathcal{E}_{\mathrm{nodiff}}$ $(-1<\beta<1)$ the boundary condition reads

$$\tau_{\mathrm{s}}\dot{U} + A_{\mathrm{b}}\tfrac{\partial u}{\partial\nu} + \omega_{\mathrm{s}}U = 0. \tag{4.26}$$

This boundary condition can be found as a special case in [Pet04].

In the case $\alpha > 1$ ($\overline{\Psi}_0 = \overline{\Psi}_{\mathrm{slow}}$) we obtain the bulk Allen-Cahn equation (4.23) with Dirichlet boundary condition, i.e., $\dot{U} = 0$. Which means that the boundary values are fixed by the initial conditions. Since we assumed in the convergence analysis that the initial energies converge, the initial values $(u(0),U(0))$ have to lie in the spaces V_{tang}, V_{const} and V_{nodiff} for $\beta = 1$, $\beta > 1$ and $-1<\beta<1$, respectively. In particular, in the first case we have $u|_\Gamma = u|_\Gamma(0)\in\mathrm{H}^1(\Gamma)$, while in the second case the boundary values are constant (on each connected component of the boundary) and in the last case we have $u|_\Gamma = u(0)|_\Gamma \in \mathrm{H}^{\frac{1}{2}}(\Gamma)$.

At last, we discuss the fast evolution case $\alpha < 1$ ($\overline{\Psi}_0 = \overline{\Psi}_{\mathrm{fast}}$). Choosing $\widetilde{\boldsymbol{u}} = \boldsymbol{u} - h\boldsymbol{w}$, $h>0$ in the limit evolutionary variational inequality (4.22) and letting $h\to 0$ we obtain the system

$$\begin{pmatrix} \tau_{\mathrm{b}}\dot{u} + \mathrm{D}_u\mathcal{E}_0(u,U) \\ \mathrm{D}_U\mathcal{E}_0(u,U)\end{pmatrix} = 0.$$

Hence, for $\beta = 1$ the limit energy functional is given by $\mathcal{E}_{\mathrm{tang}}$, and we deduce that the bulk equation (4.23) is in this case coupled to the nonlinear elliptic surface equation

$$-A_{\mathrm{s}}\Delta_\Gamma U + A_{\mathrm{b}}\tfrac{\partial u}{\partial\nu} + W_{\mathrm{s}}'(U) = 0. \tag{4.27}$$

While for $\mathcal{E}_0 = \mathcal{E}_{\text{const}}$ $(\beta > 1)$ we have the following nonlinear equation for each connected component of the boundary Γ

$$A_{\mathrm{b}}\big[\tfrac{\partial u}{\partial \nu}\big]_i + W_{\mathrm{s}}'(U^i) = 0. \tag{4.28}$$

In the last case $-1 < \beta < 1$ and therefore $\mathcal{E}_0 = \mathcal{E}_{\text{nodiff}}$ we obtain the usual Robin boundary condition

$$A_{\mathrm{b}}\tfrac{\partial u}{\partial \nu} + \omega_{\mathrm{s}} U = 0. \tag{4.29}$$

Numerical simulation

Figure 4.2 shows details of numerical simulations of equation (4.23) in a (polygonal approximation of a) circular domain using continuous piecewise affine finite elements in the bulk and on the surface (see e.g. [ElS10]). In particular, the behavior of two solutions of (4.23) near the boundary Γ is depicted for subsequent times in case of the dynamical boundary condition (4.24) and the Neumann boundary $\frac{\partial u}{\partial \nu} = 0$. Here, both solutions are starting from the same initial condition. The potential in the bulk is given by the double-well potential $W_{\mathrm{b}}(u) = \frac{k_{\mathrm{b}}}{4}(1-u^2)^2$ with $k_{\mathrm{b}} > 0$ while for the dynamic boundary condition we additionally have the quadratic potential $W_{\mathrm{s}}(U) = \frac{k_{\mathrm{s}}}{2}(1-U)^2$ on Γ with $k_{\mathrm{s}} > 0$. The dynamic boundary condition models a strong interaction between the wall and the mixture components described by the order parameter u (resp. U). In particular, due to the potential W_{s} and the surface diffusion we have an accumulation of the phase $U = 1$ at the boundary Γ , which can be clearly seen in the pictures.

4.2 An interface condition for the scalar diffusion equation

In this section we derive an interface condition for a scalar diffusion system from a bulk approximation. In particular, we consider the scalar diffusion equation in the one-dimensional domain $\Omega = \left]-\frac{1}{2}, \frac{1}{2}\right[$ given by

$$\dot{u}_\varepsilon = (a_\varepsilon(x) u_\varepsilon')' \quad \text{in } \Omega, \qquad u_\varepsilon' = 0 \ \text{ in } \left\{-\tfrac{1}{2}, +\tfrac{1}{2}\right\}, \tag{4.30}$$

where " $\dot{}$ " denotes as usual differentiation with respect to time and " $'$ " with respect to the spatial variable. Moreover, the x-dependent diffusion constant a_ε is given by the step function

$$a_\varepsilon(x) = \begin{cases} \rho & \text{for } \frac{\varepsilon}{2} \le |x| \le \frac{1}{2}, \\ \varepsilon k & \text{for } |x| < \frac{\varepsilon}{2}, \end{cases}$$

with ρ and k denoting positive constants and $\varepsilon > 0$ being a sufficiently small parameter. Let us remark here that the related stationary problem was discussed in [Att84, Sect. 1.3.5]. In particular, there the Γ-convergence of the Dirichlet energies associated with a_ε was studied in the weak topology of $\mathrm{L}^2(\Omega)$.

It was shown in Chapter 2 that the equation in (4.30) is the gradient flow of the energy functional $\mathcal{E}$ and the Wasserstein-type Onsager operator $\mathcal{K}_\varepsilon$ given respectively by

$$\mathcal{E}(u) = \int_\Omega u \log u \,\mathrm{d}x, \qquad \mathcal{K}_\varepsilon(u)\xi = -\big(a_\varepsilon(x) u\, \xi'\big)'. \tag{4.31}$$

The crucial point in the discussion of this section is that we derive the limit interface system from (4.30) using only the gradient structure given by $\mathcal{E}$ and $\mathcal{K}_\varepsilon$. This has a number of reasons: The first is that the Wasserstein formulation of diffusion equations is a natural and physically meaningful structure for this problem. In particular, the strong connection between Wasserstein gradient flows and stochastic particle systems within the theory of large deviations was shown by Peletier et al. in [AD*11, PeR11].

The second reason is that the Wasserstein gradient flow structure is known to arise in an impressively wide range of models and systems, and therefore any method that uses only the properties of this structure has the potential of application to a wide range of problems. Consequently, our approach here is to limit our use of information to those properties that follow directly from the gradient flow structure.

As a third reason, this discussion fits into a general endeavor to use gradient flow structures to pass to the limit in evolution systems using variational methods such as Γ-convergence.

We define the dissipation functional $\mathcal{J}_\varepsilon$ via

$$\mathcal{J}_\varepsilon(u,\dot u) = \frac{1}{2}\int_\Omega \Big[\frac{1}{a_\varepsilon(x)u}|w|^2 + \frac{a_\varepsilon(x)}{u}|u'|^2\Big]\,\mathrm{d}x,$$

where $w(t,x) = \int_{-1/2}^x \dot u(t,\eta)\,\mathrm{d}\eta$ denotes the primitive of $\dot u$. In particular, defining the function ξ via $\xi' = w/(a_\varepsilon u)$ we obtain $\mathcal{K}_\varepsilon(u)\xi = \dot u$. Hence, $\mathcal{J}_\varepsilon$ is nothing but the sum of the dissipation and dual dissipation potentials $\Psi_\varepsilon(u,\dot u)$ and $\Psi_\varepsilon^*(u;-\mathrm{D}\mathcal{E}(u))$, where for $\mathcal{G}_\varepsilon = \mathcal{K}_\varepsilon^{-1}$ we have $\Psi^*_\varepsilon(u;\xi) = \frac{1}{2}\langle \xi, \mathcal{K}_\varepsilon(u)\xi\rangle$ and $\Psi^*_\varepsilon(u;\xi) = \frac{1}{2}\langle \mathcal{G}_\varepsilon(u)\dot u, \dot u\rangle$, respectively.

With this a solution u_ε of (4.30) solves the equivalent formulation (see [AGS05, Lis09])

$$\forall\, t\in[0,T]: \quad \mathcal{E}(u_\varepsilon(t)) + \int_0^t \mathcal{J}_\varepsilon(u_\varepsilon(s),\dot u_\varepsilon(s))\,\mathrm{d}s = \mathcal{E}(u_\varepsilon(0)). \tag{4.32}$$

For each $\varepsilon > 0$ there exists a nonnegative solution of u_ε and it is easy to see that the quantity $\mathcal{Q}(u_\varepsilon) = \int_\Omega u_\varepsilon \,\mathrm{d}x$ is conserved along the solution u_ε. Let us fix $M > 0$ and assume that for all ε we have $\mathcal{Q}_\varepsilon(u_\varepsilon(0)) = M$.

We show in the following that u_ε converges (up to subsequences) to a limit u which satisfies the bulk equation

$$\dot u = \rho u'' \quad \text{in } \big]{-\tfrac{1}{2}},0\big[\;\cup\; \big]0,\tfrac{1}{2}\big[, \tag{4.33a}$$

coupled by the interface condition

$$\rho u'_- = k(u_+ - u_-) = \rho u'_+ \qquad \text{in } \{0\}, \tag{4.33b}$$

where the indices $_+$ and $_-$ denote the limits from the right and from the left, respectively.

As was shown in [GlM13] (see also [Mie13]) we can write this limit system as the gradient system

$$
\begin{aligned}
\mathcal{E}(u) &= \int_\Omega u \log u \,\mathrm{d}x, \quad \text{and} \\
\Psi_0^*(u,\xi) &= \frac{1}{2}\int_{\Omega\setminus\{0\}} \rho u |\xi'|^2 \,\mathrm{d}x + \frac{k}{2}\Lambda(u_+,u_-)(\xi_+-\xi_-)^2,
\end{aligned}
\tag{4.34}
$$

where as usual $\Lambda(a,b) = (a{-}b)/(\log a{-}\log b)$ denotes the logarithmic mean of a and b.

In the theory of semiconductor devices interface conditions like (4.33b) are known as *thermionic emission* and model the transport of electrons or holes over a heterointerface, see [Sch94, SzN07]. Notably, nonstandard interface and transmission conditions in semiconductor heterostructures and biological systems are of great importance (see [Gli12] and [ElS10]). Especially in organic photovoltaics interfaces are the fundamental building block, see [PoA06, Section 8].

4.2.1 Transformation of the domain

The first step in the convergence proof of the system (4.32) is a rescaling of the domain Ω which stretches the region around $x = 0$. This converts the functions u_ε, which have steep gradients around $x = 0$ into functions U_ε that will have a more regular behavior.

However, let us remark here that for the derivation of interface conditions in higher dimensions, i.e., $\Omega \subset \mathbb{R}^d$ with $d \geq 2$, and non-trivial geometry of the interface a direct argument without the detour of rescaling is preferable. Otherwise one has to resort to the quite technical approach of flattening the interface using a partition of unity and local coordinates.

We introduce the piecewise affine transformation $x \mapsto y_\varepsilon(x) \in \,]-1,+1[$

$$
y_\varepsilon(x) = \begin{cases} \frac{x}{\varepsilon} & \text{for } |x| \leq \frac{\varepsilon}{2}, \\ \frac{x \pm c_\varepsilon}{1-\varepsilon} & \text{for } \frac{\varepsilon}{2} < |x| < \frac{1}{2}, \end{cases}
$$

where $c_\varepsilon = (1{-}2\varepsilon)/2$. This stretches the shrinking interval $]-\frac{\varepsilon}{2},\frac{\varepsilon}{2}[$ to the fixed interval $]-\frac{1}{2},\frac{1}{2}[$. Now, with a function $u:\Omega\to\mathbb{R}$ we associate a function U defined on $\Sigma \overset{\text{def}}{=} \,]-1,+1[$ via $u(x) = U(y_\varepsilon(x))$ such that

$$
u'(x) = \begin{cases} \frac{1}{\varepsilon}\, U'(y_\varepsilon(x)) & \text{for } |x| \leq \frac{\varepsilon}{2}, \\ \frac{1}{1-\varepsilon}\, U'(y_\varepsilon(x)) & \text{for } \frac{\varepsilon}{2} < |x| < \frac{1}{2}. \end{cases}
$$

We can easily transport the functionals $\mathcal{E}$ and $\mathcal{J}_\varepsilon$ to the new setting by defining

$$
\begin{aligned}
\overline{\mathcal{E}}_\varepsilon(U) &= \int_\Sigma \gamma_\varepsilon(y) U \log U \,\mathrm{d}y, \qquad \text{and} \\
\overline{\mathcal{J}}_\varepsilon(U,\dot U) &= \frac{1}{2}\int_\Sigma \Big[\frac{1}{\alpha_\varepsilon(y)\,U}|W|^2 + \frac{\alpha_\varepsilon(y)}{U}|U'|^2\Big]\,\mathrm{d}y,
\end{aligned}
$$

where $W(t,y) = \int_{-1}^{y} \gamma_\varepsilon(\eta)\dot U(t,\eta)\,\mathrm{d}\eta$ and the step functions α_ε and γ_ε are defined as

$$(\alpha_\varepsilon(y),\gamma_\varepsilon(y)) = \begin{cases} (k,\varepsilon) & \text{for } |y| \le \frac12, \\ (\frac{\rho}{1-\varepsilon}, 1-\varepsilon) & \text{for } \frac12 < |y| < 1. \end{cases}$$

Hence, letting U_ε denote the rescaled function associated with the solution u_ε we find that U_ε satisfies the corresponding formulation

$$\overline{\mathcal{E}}_\varepsilon(U_\varepsilon(t)) + \int_0^t \overline{\mathcal{J}}_\varepsilon(U_\varepsilon(s),\dot U_\varepsilon(s))\,\mathrm{d}s = \overline{\mathcal{E}}_\varepsilon(U_\varepsilon(0)). \tag{4.35}$$

Moreover, $\overline{\mathcal{Q}}_\varepsilon(U_\varepsilon) = \int_\Sigma \gamma_\varepsilon(y)U_\varepsilon(y)\,\mathrm{d}y = \mathcal{Q}(u_\varepsilon)$ is conserved, i.e.

$$\overline{\mathcal{Q}}_\varepsilon(U_\varepsilon) = M \qquad \text{such that} \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\overline{\mathcal{Q}}_\varepsilon(U_\varepsilon) = \overline{\mathcal{Q}}_\varepsilon(\dot U_\varepsilon) = 0. \tag{4.36}$$

Note that we can safely assume that u_ε is uniformly bounded in $\mathrm{L}^\infty(\Omega_T)$, where $\Omega_T = \Omega\times[0,T]$, and thus we also have that

$$U_\varepsilon \text{ is uniformly bounded in } \mathrm{L}^\infty(\Sigma_T). \tag{4.37}$$

4.2.2 Passing to the limit

Let us first start by deriving compactness properties of the sequence U_ε. In particular, let us additionally make the reasonable assumption that the initial energies $\overline{\mathcal{E}}_\varepsilon(U_\varepsilon(0))$ are bounded. Then, by (4.35), we have the boundedness of $\int_0^T \overline{\mathcal{J}}_\varepsilon(U_\varepsilon,\dot U_\varepsilon)\,\mathrm{d}t$ which in turn implies that there exists a constant $C>0$ such that

$$\int_0^T\int_\Sigma \frac{|W_\varepsilon|^2}{U_\varepsilon}\,\mathrm{d}y\,\mathrm{d}t < C, \quad \text{and} \quad \int_0^T\int_\Sigma \frac{|U_\varepsilon'|^2}{U_\varepsilon}\,\mathrm{d}y\,\mathrm{d}t < C,$$

where $W_\varepsilon(t,y) \overset{\mathrm{def}}{=} \mathcal{I}_\varepsilon[\dot U_\varepsilon](t,y)$. Since the sequence U_ε also happens to be bounded in $\mathrm{L}^\infty(\Sigma_T)$ the bounds above imply weak compactness of W_ε and U_ε' in $\mathrm{L}^2(\Sigma_T)$. In particular, we can extract a (not relabeled) subsequence such that

$$U_\varepsilon \overset{*}{\rightharpoonup} U \text{ in } \mathrm{L}^\infty(\Sigma_T), \qquad U_\varepsilon' \rightharpoonup U' \quad \text{and} \quad W_\varepsilon \rightharpoonup W \text{ in } \mathrm{L}^2(\Sigma_T). \tag{4.38}$$

Here, the limit W has to be determined. We will show below that W corresponds to the distributional time derivative of U.

The following lemma establishes a lower estimate for the dissipation functional $\overline{\mathcal{J}}_\varepsilon$.

Lemma 4.2.1 *For W,U in* (4.38) *and for all $t\in[0,T]$ it holds that*

$$\liminf_{\varepsilon\to0}\int_0^t \overline{\mathcal{J}}_\varepsilon(U_\varepsilon,\dot U_\varepsilon)\,\mathrm{d}s \ge \int_0^t\int_\Sigma \left[\frac{1}{\alpha(y)U}|W|^2 + \frac{\alpha(y)}{U}|U'|^2\right]\mathrm{d}y\,\mathrm{d}s, \tag{4.39}$$

where $\alpha(y)=\rho$ for $\frac12 \le |y| \le 1$ and $\alpha(y)=k$ for $|y|<\frac12$.

Proof: For arbitrary $A \in \mathrm{L}^1(\Sigma_T)$, $B \in \mathrm{L}^2(\Sigma_T)$ satisfying $A + B^2/2 \le 0$ almost everywhere in Σ_T we have

$$\begin{aligned}\liminf_{\varepsilon\to 0} \frac{1}{2}\int_0^t \int_\Sigma \frac{|W_\varepsilon|^2}{\alpha_\varepsilon U_\varepsilon}\,\mathrm{d}y\,\mathrm{d}s &\ge \liminf_{\varepsilon\to 0}\int_0^t\int_\Sigma \frac{U_\varepsilon}{\alpha_\varepsilon}\left[A + \frac{B^2}{2} + \frac{1}{2}\left(\frac{W_\varepsilon}{U_\varepsilon}\right)^2\right]\mathrm{d}y\,\mathrm{d}s \\ &\ge \liminf_{\varepsilon\to 0}\int_0^t\int_\Sigma \frac{AU_\varepsilon + BW_\varepsilon}{\alpha_\varepsilon}\,\mathrm{d}y\,\mathrm{d}s = \int_0^t\int_\Sigma \frac{AU + BW}{\alpha}\,\mathrm{d}y\,\mathrm{d}s.\end{aligned}$$

Now, choosing sequences A_n, B_n that approximate $-|W/U|^2/2$ and W/U, respectively, and using Fatou's Lemma we arrive at

$$\liminf_{\varepsilon\to 0}\frac{1}{2}\int_0^t\int_\Sigma \frac{|W_\varepsilon|^2}{\alpha_\varepsilon U_\varepsilon}\,\mathrm{d}s \ge \frac{1}{2}\int_0^t\int_\Sigma \frac{|W|^2}{\alpha U}\,\mathrm{d}s.$$

Arguing analogously for $|U_\varepsilon'|^2/U_\varepsilon$ we finally obtain the lower estimate in (4.39). □

We define the limit energy $\overline{\mathcal{E}}_0$ via

$$\overline{\mathcal{E}}_0(U) = \int_{-1}^{-1/2} U\log U\,\mathrm{d}y + \int_{1/2}^{1} U\log U\,\mathrm{d}y.$$

In the following lemma we prove that $\overline{\mathcal{E}}_0$ is a lower limit for $\overline{\mathcal{E}}_\varepsilon$.

Lemma 4.2.2 *For U as in* (4.38) *the liminf estimate*

$$\liminf_{\varepsilon\to 0}\overline{\mathcal{E}}_\varepsilon(U_\varepsilon(t)) \ge \overline{\mathcal{E}}_0(U(t)) \tag{4.40}$$

holds for all $t \in [0,T]$.

Proof: First of all let us note that due to $\gamma_\varepsilon \dot U_\varepsilon = W_\varepsilon'$ we have for all $0 \le t_0 < t_1 \le T$ and $\Phi \in \mathrm{C}^1(\overline{\Sigma})$

$$\int_\Sigma \gamma_\varepsilon U_\varepsilon(t_1)\Phi\,\mathrm{d}y - \int_\Sigma \gamma_\varepsilon U_\varepsilon(t_0)\Phi\,\mathrm{d}y = -\int_{t_0}^{t_1}\int_\Sigma W_\varepsilon \Phi'\,\mathrm{d}y\,\mathrm{d}t.$$

(Note that due to its definition W_ε vanishes for $y = \pm 1$.) Hence, we can argue as in [AM*12, Proof of Theorem 3.1] by recalling the definition of the 1-Wasserstein distance, denoted $\mathsf{d}_{\mathrm{W}_1}$, for the family of measures $\mathrm{d}\rho_\varepsilon(t) = \gamma_\varepsilon U_\varepsilon(t)\,\mathrm{d}y$ (see [AGS05, Theorem 6.1.1])

$$\begin{aligned}\mathsf{d}_{\mathrm{W}_1}(\rho_\varepsilon(t_1),\rho_\varepsilon(t_0)) &\overset{\mathrm{def}}{=} \sup\left\{\int_\Sigma \Phi\,\mathrm{d}\rho_\varepsilon(t_1) - \int_\Sigma \Phi\,\mathrm{d}\rho_\varepsilon(t_0) \;:\; \Phi \in \mathrm{C}^1(\overline{\Sigma}),\ |\Phi'| \le 1\right\} \\ &\le \int_{t_0}^{t_1}\int_\Sigma |W_\varepsilon|\,\mathrm{d}y\,\mathrm{d}t \le \sqrt{2(t_1-t_0)}\left(\int_0^T\int_\Sigma |W_\varepsilon|^2\,\mathrm{d}y\,\mathrm{d}t\right)^{1/2}.\end{aligned}$$

It follows by the boundedness of W_ε that the curves $t \mapsto \rho_\varepsilon(t)$ are an equicontinuous family of mappings from $[0,T]$ into the space $\mathsf{Meas}(\Sigma)$ endowed with the 1-Wasserstein distance. Since we have the total mass $\rho_\varepsilon(t,\Sigma) = M = \overline{\mathcal{Q}}_\varepsilon(U_\varepsilon)$, we can apply the generalized Arzelà-Ascoli theorem [AGS05, Proposition 3.3.1] to obtain a (not relabeled) subsequence such that for all $t \in [0,T]$ and all $\Phi \in \mathrm{C}_b(\Sigma)$

$$\int_\Sigma \Phi \,\mathrm{d}\rho_\varepsilon(t) \quad \longrightarrow \quad \int_{-1}^{-1/2} U(t)\Phi \,\mathrm{d}y + \int_{1/2}^{1} U(t)\Phi \,\mathrm{d}y. \tag{4.41}$$

Moreover, since the limit is uniquely characterized by the weak*-convergence of U_ε in (4.38) the whole sequence converges.

Next, observe that for all $t \in [0,T]$ we have the identity

$$\frac{1}{2}\int_\Sigma \gamma_\varepsilon |U_\varepsilon(t)|^2 \,\mathrm{d}y = \frac{1}{2}\int_\Sigma \gamma_\varepsilon |U_\varepsilon(0)|^2 \,\mathrm{d}y - \int_0^t \int_\Sigma W_\varepsilon U_\varepsilon' \,\mathrm{d}y \,\mathrm{d}s.$$

In particular, due to the boundedness of W_ε and U_ε' the right-hand side above is bounded and hence also $\int_{-1/2}^{1/2} \varepsilon |U_\varepsilon(t)|^2 \,\mathrm{d}y$ for all $t \in [0,T]$. Thus, for all $\rho > 0$ we have that $\varepsilon^{1/2+\rho} U_\varepsilon(t) \to 0$ strongly in $\mathrm{L}^2\left(\left[-\frac{1}{2}, \frac{1}{2}\right]\right)$ and we infer

$$\forall\, t \in [0,T]: \qquad \int_{-1/2}^{1/2} \varepsilon U_\varepsilon(t) \log U_\varepsilon(t) \,\mathrm{d}y \longrightarrow 0.$$

Using this and (4.41) the lower semicontinuity properties of the free energy (see [AGS05, Lemma 9.4.3] or [ASZ09, Lemma 6.2]) yield the liminf estimate. □

Remark 4.2.3 *Following the argumentation in the proof of Lemma 4.2.2 we also infer that the limit U satisfies $\int_{-1}^{-1/2} U(t,y)\,\mathrm{d}y + \int_{1/2}^{1} U(t,y)\,\mathrm{d}y = M$.*

In order to characterize the limit system we identify the limit function W first: Using $W_\varepsilon(t,-1) = W_\varepsilon(t,1) = 0$ we have for $\Phi \in \mathrm{C}^1([0,T] \times \Sigma)$

$$\int_s^t \int_\Sigma W_\varepsilon \Phi' \,\mathrm{d}y \,\mathrm{d}r = \int_s^t \int_\Sigma \gamma_\varepsilon U_\varepsilon \dot{\Phi} \,\mathrm{d}y \,\mathrm{d}r + \int_\Sigma \gamma_\varepsilon \big[U_\varepsilon(s)\Phi(s) - U_\varepsilon(t)\Phi(t)\big] \,\mathrm{d}y.$$

Thus, using the convergence of W_ε and $\gamma_\varepsilon U_\varepsilon$ we can pass to the limit for $\varepsilon \to 0$ to obtain

$$\int_s^t \int_\Sigma W \Phi' \,\mathrm{d}y \,\mathrm{d}r = \int_s^t \int_{\Sigma_0} U \dot{\Phi} \,\mathrm{d}y \,\mathrm{d}r + \int_{\Sigma_0} \big[U(s)\Phi(s) - U(t)\Phi(t)\big] \,\mathrm{d}y, \tag{4.42}$$

where $\Sigma_0 = \left]-1, -\frac{1}{2}\right[\cup \left]\frac{1}{2}, 1\right[$ denotes the “bulk” domain. Hence, the limit W corresponds to the distributional time derivative $\dot{U} \in \mathrm{L}^2(0,T;\mathrm{H}^1(\Sigma_0)^*)$. Moreover, choosing Φ such that $\operatorname{supp}\Phi \subset \Sigma\backslash\Sigma_0$ we find that

$$\text{for fixed } t \in [0,T] \quad W(t,\cdot) \quad \text{is constant in} \quad \Sigma\backslash\Sigma_0. \tag{4.43}$$

In the following lemma we establish the chain rule for $t \mapsto \overline{\mathcal{E}}_0(U(t))$.

Lemma 4.2.4 (Chain rule) *For all* $0 \le s < t \le T$ *it holds that*

$$\int_s^t \int_\Sigma \frac{WU'}{U} \,\mathrm{d}y\,\mathrm{d}r = \overline{\mathcal{E}}_0(U(s)) - \overline{\mathcal{E}}_0(U(t)). \tag{4.44}$$

Proof: First, note that the left-hand side in (4.44) is well-defined due to Lemma 4.2.1 and Hölder's inequality. We argue as in [Gli12] and define the "truncated" energy functionals $\overline{\mathcal{E}}_0^\rho$ via

$$\overline{\mathcal{E}}_0^\rho(U) = \int_{\Sigma_0} \int_0^{U(y)} [1 + \log(\max\{\rho, w\})] \,\mathrm{d}w\,\mathrm{d}y.$$

We easily check that $\overline{\mathcal{E}}_0^\rho(U) \to \overline{\mathcal{E}}_0(U)$ holds for $\rho \downarrow 0$. Moreover, we apply [Bré73, Lemma 3.3] to obtain the chain rule for $t \mapsto \overline{\mathcal{E}}_0^\rho(U(t))$, i.e.,

$$\int_s^t \int_\Sigma \frac{WU'_\rho}{U_\rho} \,\mathrm{d}y\,\mathrm{d}r = \overline{\mathcal{E}}_0^\rho(U(s)) - \overline{\mathcal{E}}_0^\rho(U(t)),$$

where we used the notation $U_\rho = \max\{U, \rho\}$ and that we have $(\log U_\rho)' = U'_\rho / U_\rho$. Now, using the dominated convergence theorem we let $\rho \downarrow 0$ and arrive at (4.44). $\square$

In the final step of our limit passage we transfer all quantities back to the domain $\Omega \backslash \{0\}$ to characterize the limit system. In particular, we define the function $u : [0, T] \times (\Omega \backslash \{0\}) \to \mathbb{R}$ via

$$u(t,x) = \begin{cases} U(t, x - \frac{1}{2}) & \text{for } x \in \left]-\frac{1}{2}, 0\right[, \\ U(t, x + \frac{1}{2}) & \text{for } x \in \left]0, \frac{1}{2}\right[. \end{cases} \tag{4.45}$$

Theorem 4.2.5 *The function* u *defined in* (4.45) *is a solution of the limit system* (4.33)*. In particular,* u *satisfies*

$$\forall\, t \in [0, T]: \qquad \mathcal{E}(u(t)) + \int_0^t \Psi_0(u, \dot{u}) + \Psi_0^*(u, -\mathrm{D}\mathcal{E}(u)) \,\mathrm{d}s = \mathcal{E}(u(0)).$$

Proof: Combining (4.39) and (4.40) and assuming additionally that the initial energies satisfy $\overline{\mathcal{E}}_\varepsilon(U_\varepsilon(0)) \to \overline{\mathcal{E}}_0(U(0))$ we get the lower estimate

$$\overline{\mathcal{E}}_0(U(t)) + \frac{1}{2} \int_0^t \int_\Sigma \left[\frac{1}{\alpha U} |W|^2 + \frac{\alpha}{U} |U'|^2 \right] \mathrm{d}y\,\mathrm{d}s \le \overline{\mathcal{E}}_0(U(0)). \tag{4.46}$$

Let us denote by $\overline{\mathcal{M}}(W, U)$ the inner integral. Using the binomial formula we find

$$\int_0^t \overline{\mathcal{M}}(W, U) \,\mathrm{d}s = \int_0^t \int_\Sigma \left[\frac{(W - \alpha U')^2}{2\alpha U} + \frac{W\,U'}{U} \right] \mathrm{d}y\,\mathrm{d}s. \tag{4.47}$$

Using the chain rule in Lemma 4.2.4 we infer that the lower energy estimate (4.46) is equivalent to

$$\forall\, t \in [0,T]: \qquad \int_0^t \int_\Sigma \frac{\big(W - \alpha U'\big)^2}{\alpha U} \,\mathrm{d}y \,\mathrm{d}s = 0.$$

Thus, we have that $W = \alpha U'$ almost everywhere in $[0,T] \times \Sigma$. In particular, since W is constant in $\Sigma \backslash \Sigma_0$ we have that U is affine such that for $U_\pm = U(\pm\frac{1}{2})$

$$y \mapsto U(y) = (U_+ {-} U_-) y + \frac{U_+ {+} U_-}{2} \qquad \text{for } y \in \Sigma \backslash \Sigma_0.$$

Moreover, using (4.42) and the definition of u in (4.45) we get for all $0 \le s < t \le T$ and $\Phi \in \mathrm{C}^1(0,T;\mathrm{H}^1(\Sigma_0))$

$$\begin{aligned} 0 &= \int_s^t \int_{\Sigma_0} \big[U \dot{\Phi} - \rho U' \Phi'\big] \,\mathrm{d}y \,\mathrm{d}r + \int_s^t k (U_+ {-} U_-)(\Phi_+ {-} \Phi_-) \,\mathrm{d}r \\ &\quad + \int_{\Sigma_0} \big[U(s)\Phi(s) - U(t)\Phi(t)\big] \,\mathrm{d}y \\ &= \int_s^t \int_{\Omega\backslash\{0\}} \big[u \dot{\phi} - \rho u' \phi'\big] \,\mathrm{d}x \,\mathrm{d}r + \int_s^t k (u_+ {-} u_-)(\phi_+ {-} \phi_-) \,\mathrm{d}r \\ &\quad + \int_{\Omega\backslash\{0\}} \big[u(s)\phi(s) - u(t)\phi(t)\big] \,\mathrm{d}x, \end{aligned}$$

where ϕ is associated with Φ as in (4.45). Thus, u is indeed a weak solution of the limit system (4.33).

Let us remark that we have the identity $\overline{\mathcal{E}}_0(U) = \mathcal{E}(u)$. Moreover, we define the function $(t,y) \mapsto \Xi(t,y)$ via $\Xi' = W/(\alpha U)$. Since W is constant (for fixed t) and U is affine in $\Sigma \backslash \Sigma_0$ we obtain by integration

$$\Xi_+ - \Xi_- = \frac{1}{k} \int_{-1/2}^{1/2} \frac{W}{U(\eta)} \,\mathrm{d}\eta = \frac{W}{k} \frac{\log U_+ - \log U_-}{U_+ - U_-}.$$

In particular, we have $W = k \Lambda(u_+, u_-)(\xi_+ {-} \xi_-)$ for ξ associated with Ξ, and we obtain the limit dissipation potential

$$\frac{1}{2} \int_\Sigma \frac{|W|^2}{\alpha U} \,\mathrm{d}y = \frac{1}{2} \int_{\Omega\backslash\{0\}} \rho u |\xi'|^2 \,\mathrm{d}y + \frac{k}{2} \Lambda(u_+, u_-) \big(\xi_+ {-} \xi_-\big)^2 = \Psi_0(u; \dot{u}).$$

Finally, we directly compute

$$\frac{1}{2} \int_\Sigma \frac{\alpha |U'|^2}{U} \,\mathrm{d}y = \int_{\Omega\backslash\{0\}} \frac{\rho |u'|^2}{u} \,\mathrm{d}x + \frac{k}{2} (u_+ {-} u_-) \big(\log u_+ {-} \log u_-\big) = \Psi_0^*(u, -\mathrm{D}\mathcal{E}(u))$$

and recover the dual dissipation potential of the limit system. $\square$

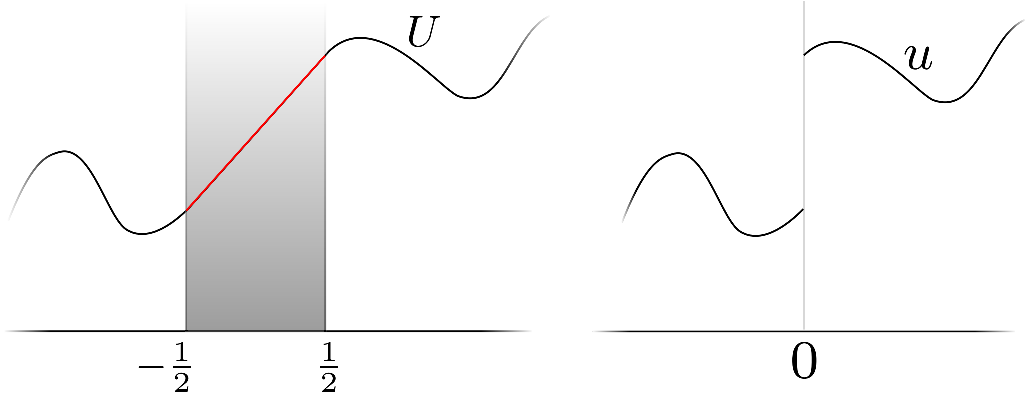

Figure 4.3: Left: Sketch of the limit U, which is affine in the stretched interface layer $[-\frac{1}{2}, \frac{1}{2}]$ (gray). Right: The associated function u in (4.45) with jump across the interface $\{0\}$.

4.2.3 Geodesic λ-convexity of the interface system

In this final section we comment on the applicability of the methods developed in Chapter 3 to the system $\mathcal{E}, \mathcal{K}_\varepsilon$ in (4.31). In particular, we are interested in the possibility of using the evolutionary variational inequality (EVI) formulation to derive the limit system in (4.33). Since the theory of Chapter 3 works only for $a_\varepsilon \in \mathrm{W}^{2,\infty}(\overline{\Omega})$ let us consider a smooth mollification of a_ε, which we will denote by $a_{\varepsilon\rho}$. Moreover, as in Section 3.2.1 we introduce the distance $\mathsf{d}_{\varepsilon\rho}$ induced by the Onsager operator $\mathcal{K}_{\varepsilon\rho}$, where $\mathcal{K}_{\varepsilon\rho}(u)\xi = -(a_{\varepsilon\rho}(x)u\xi')'$. Formula (3.45) on Page 35 then shows that $\mathcal{E}$ is geodesically $\lambda_{\varepsilon\rho}$-convex with respect to $\mathsf{d}_{\varepsilon\rho}$ with

$$\lambda_{\varepsilon\rho} = \inf\left\{-a''_{\varepsilon\rho}(x)/2 + (a'_{\varepsilon\rho}(x))^2/(4a_{\varepsilon\rho}(x)) \,:\, x \in \Omega\right\}.$$

Hence, the solution $u_{\varepsilon\rho}$ also satisfies the evolutionary variational inequality

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\mathsf{d}_{\varepsilon\rho}(u_{\varepsilon\rho}(t), v)^2 + \frac{\lambda_{\varepsilon\rho}}{2}\mathsf{d}_{\varepsilon\rho}(u_{\varepsilon\rho}(t), v)^2 + \mathcal{E}(u_{\varepsilon\rho}(t)) \leq \mathcal{E}(v).$$

For a suitable $\rho = \rho(\varepsilon)$, with $\rho(\varepsilon) \to 0$ when $\varepsilon \to 0$, we aim at passing to the limit $\varepsilon \to 0$. However, some simple considerations show that $\lambda_{\varepsilon\rho(\varepsilon)} \to -\infty$ for any choice of $\rho(\varepsilon)$. Hence, we cannot exploit the geodesic λ-convexity of $\mathcal{E}$ in this particular case.

Moreover, if we directly apply the machinery developed in Chapter 3 to the limit system $(\mathcal{E}, \Psi_0)$ we also obtain a negative result: The corresponding limit Onsager operator $\mathcal{K}_0$ is defined as

$$\mathcal{K}_0(u)\xi = -(\delta u\xi')', \quad \text{for } \xi \text{ satisfying} \quad \delta u_+\xi'_+ = \delta u_-\xi'_- = k\Lambda(u_+, u_-)(\xi_+ - \xi_-).$$

We compute the form $\langle \xi, \mathcal{M}_0(u)\xi\rangle = \langle \xi, \mathrm{D}\mathcal{F}_0(u)\mathcal{K}_0(u)\xi\rangle - \frac{1}{2}\langle \xi, \mathrm{D}\mathcal{K}_0(u)[\mathcal{F}_0(u)]\xi\rangle$, where

$\mathcal{F}_0(u) = \mathcal{K}_0(u)\mathrm{D}\mathcal{E}(u)$ denotes as usual the vector field. In particular, we have

$$\begin{aligned} \mathcal{F}_0(u) &= -\delta u'' \quad \text{for } u \text{ satisfying} \quad \delta u'_+ = \delta u'_- = k(u_+ - u_-), \\ \mathrm{D}\mathcal{F}_0(u)[v] &= -\delta v'' \quad \text{for } v \text{ satisfying} \quad \delta v'_+ = \delta v'_- = k(v_+ - v_-). \end{aligned}$$

Hence, denoting $v = -(\delta u \xi')'$ we obtain

$$\begin{aligned} \langle \xi, \mathcal{M}_0(u)\xi \rangle = &\int_{\Omega\setminus\{0\}} -\delta \xi v'' \,\mathrm{d}x + \frac{1}{2}\int_{\Omega\setminus\{0\}} \delta^2 u'' |\xi'|^2 \,\mathrm{d}x \\ &+ \frac{k\delta}{2}\Big(\partial_a \Lambda(u_+, u_-) u''_+ + \partial_b \Lambda(u_+, u_-) u''_-\Big)(\xi_+ - \xi_-)^2. \end{aligned}$$

Denoting the last term by δ and using integration by parts several times we get

$$\begin{aligned} \langle \xi, \mathcal{M}_0(u)\xi \rangle = &\int_{\Omega\setminus\{0\}} \delta^2 \xi''(u\xi')' \,\mathrm{d}x + \int_{\Omega\setminus\{0\}} \delta^2 u(\xi'''\xi' + |\xi''|^2) \,\mathrm{d}x + \delta \\ &+ \delta\big(\xi_+ v'_+ - \xi_- v'_- + v_- \xi'_- - v_+ \xi'_+\big) + \frac{\delta^2}{2}\Big(u'_- |\xi'_-|^2 - u'_+ |\xi'_+|^2 - 2u_- \xi''_- \xi'_- + 2u_+ \xi''_+ \xi'_+\Big). \end{aligned}$$

Finally, integrating by parts one last time yields

$$\begin{aligned} \langle \xi, \mathcal{M}_0(u)\xi \rangle = &\int_{\Omega\setminus\{0\}} \delta^2 u |\xi''|^2 \,\mathrm{d}x + \frac{k\delta}{2}\Big(\partial_a \Lambda(u_+, u_-) u''_+ + \partial_b \Lambda(u_+, u_-) u''_-\Big)(\xi_+ - \xi_-)^2 \\ &+ \delta\big(\xi_+ v'_+ - \xi_- v'_- + v_- \xi'_- - v_+ \xi'_+\big) + \frac{\delta^2}{2}\Big(u'_- |\xi'_-|^2 - u'_+ |\xi'_+|^2\Big). \end{aligned}$$

Although we can employ the interface conditions for ξ, v and u it is easy to check that we are not able to proceed in the computations since we cannot treat the second order terms (which have no sign) at the interface using integration by parts (e.g. see the linear reaction-diffusion system in 3.3.5).

Hence, we conjecture that the limit system is *not* geodesically λ-convex for any $\lambda \in \mathbb{R}$.

Part II

The Weighted Inertia-Dissipation-Energy principle

5 Introduction to Part II

From a general scientific viewpoint the investigation of variational principles is of a paramount importance for it corresponds to the fundamental quest for *general* and *simple* explanations of reality as we experience it. On the other hand, beside their indisputable elegance, variational principles have a clear practical impact as they originate a wealth of new perspectives and serve as unique tools for the analysis of real physical situations. Correspondingly, the mathematical literature on variational principles in mechanics is overwhelming and a number of monographs on the subject are available. Being completely beyond our purposes to attempt a comprehensive review of the development of this subject, we shall minimally refer to the classical monographs by Lánczos [Lán70] and Moiseiwitsch [Moi04] as well as to the more recent ones by Basdevant [Bas07], Berdichevsky [Ber09] and Ghoussoub [Gho09].

In Part II we present the results of [LiS13a] and [LiS13b]. In these two articles a new variational principle for semilinear partial differential equations of the form

$$\rho u'' + \nu u' - \operatorname{div}(\mathbb{A}\nabla u) + f(u) = 0 \quad \text{in } \Omega \times]0,T[\tag{5.1}$$

is discussed. Here, $\Omega \subset \mathbb{R}^d$ is a bounded and smooth domain and $T \in]0,\infty]$ is some reference time. Note that we admit the case $T = \infty$. The *density* ρ and the *viscosity* ν are nonnegative parameters satisfying $\rho + \nu > 0$. For $\rho > 0$ equation (5.1) is the (weakly damped) wave equation, but our discussion includes the limiting cases of the semilinear wave equation ($\nu = 0$) and the semilinear heat equation ($\rho = 0$) as well. For simplicity we complement equation (5.1) with homogeneous Dirichlet boundary conditions on $\partial\Omega$ and initial conditions $u(0) = u^0$, $\rho u'(0) = \rho u^1$. Here we used ρ in the fixing of the initial conditions to emphasize that we include the case $\rho = 0$ in which $u(0) = u^0$ is the only initial condition. The literature on the semilinear wave equation (5.1) is vast and it is clearly beyond the purposes of this text to provide a comprehensive review. The reader is referred to the monographs by Lions [Lio69], Shatah & Struwe [ShS98] and Lax [Lax06] for a collection of results, references, and historical remarks.

The aim of this part is to reformulate the evolutionary problem in (5.1) in a variational form. This reformulation is accomplished by introducing a functional whose minimizers represent *entire trajectories* of the system. In particular, for all $\varepsilon > 0$ we shall be concerned with the functional

$$\begin{aligned} \boldsymbol{W}_\varepsilon[u] &= \int_0^T \mathrm{e}^{-t/\varepsilon} \left[\frac{\varepsilon^2\rho}{2}\|u''\|^2 + \frac{\varepsilon\nu}{2}\|u'\|^2 + \mathcal{E}(u) \right] \mathrm{d}t, \\ \text{with} \quad \mathcal{E}(u) &= \int_\Omega \left[\tfrac{1}{2}\nabla u \cdot \mathbb{A}\nabla u + F(u) \right] \mathrm{d}x. \end{aligned} \tag{5.2}$$

The functional $\boldsymbol{W}_\varepsilon$, being defined on a suitable space $\mathbb{Y}$ of trajectories, is called *Weighted Inertia-Dissipation-Energy* functional (abbreviated WIDE functional in the following) as it features the *weighted* sum of the *inertial* term $\rho\|u''\|^2/2$, the *dissipative* term $\nu\|u'\|^2/2$, and the *energetic* term $\mathcal{E}(u)$. Note that the small parameter ε has the physical dimension of time, so that the whole integrand in $\boldsymbol{W}_\varepsilon$ is an energy and $\boldsymbol{W}_\varepsilon$ is an action. See Section 6.1 below for a formal derivation of the functional $\boldsymbol{W}_\varepsilon$ by means of time-discretized incremental problems.

Under conditions of sufficient smoothness, the Euler-Lagrange equations of the WIDE functional read

$$\varepsilon^2\rho u'''' - 2\varepsilon\rho u''' + (\rho-\varepsilon\nu)u'' + \nu u' + \mathrm{D}\mathcal{E}(u) = 0. \tag{5.3}$$

In the case of a finite time horizon $T < \infty$ the equation is complemented by the following initial *and* final conditions (see also Section 6.1 for a discussion of other final conditions)

$$\begin{gathered} u(0) = u^0, \quad \rho u'(0) = \rho u^1 \quad \text{and} \\ \varepsilon^2\rho u''(T) = 0, \quad \varepsilon^2\rho u'''(T) = \varepsilon\nu u'(T). \end{gathered} \tag{5.4}$$

In case of the infinite time horizon we will see (Section 7.2) that we have integrability conditions instead of the final conditions above. We will discuss each of the cases – finite and infinite time horizon – separately in Chapters 6 and 7, respectively.

The minimization of $\boldsymbol{W}_\varepsilon$ corresponds to an *elliptic regularization* in time of the original problem (5.1). Hence, (5.1) is replaced by the minimum problem

$$\inf_{u\in\mathbb{K}(u^0,u^1)} \boldsymbol{W}_\varepsilon(u), \tag{5.5}$$

where $\mathbb{K}(u^0, u^1) \subset \mathbb{Y}$ denotes a (affine) subspace, which encodes the initial conditions.

The crucial question is whether minimizers u_ε of $\boldsymbol{W}_\varepsilon$ (provided they exist) converge in a certain sense to a limit u which solves the original problem, i.e.,

$$\lim_{\varepsilon\to 0} u_\varepsilon = u \quad \text{solves (5.1)}. \tag{5.6}$$

The interest of this perspective resides in the possibility of connecting the *difficult* semilinear PDE problem (5.1) with a comparably *easier* problem: the constrained minimization of the functional $\boldsymbol{W}_\varepsilon$. This possibility provides a novel variational insight to the differential problem by opening the way to the application of the tools of the calculus of variations to (5.1). For instance, under certain assumptions on the functional $\boldsymbol{W}_\varepsilon$ we are able to show uniform convexity, thus it admits a unique minimizer whereas no uniqueness is known for (5.1) under general nonlinearities F'. In this regard, the WIDE functional approach can be expected to possibly serve as a *variational selection criterion* in some non-uniqueness situation (see Section 7.3 for an ODE example).

Clearly equation (5.1) is nothing but the formal limit in (5.3) for $\varepsilon \to 0$. Note that, as the above problem is of fourth order in time, the two extra *final conditions* (resp. integrability conditions) arise and, at all levels $\varepsilon > 0$, causality is lost. Owing to this fact, the convergence (5.6) is generally referred to as the *causal limit* for it results in restoring causality.

Our interest in WIDE functionals has been inspired by a conjecture by De Giorgi [De 96] on hyperbolic evolution. In particular, in [De 96] it is conjectured that the minimizers of the functional

$$u \mapsto \int_0^\infty \int_{\mathbb{R}^d} \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2}{2}|u''|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{p}|u|^p \right) \mathrm{d}x \ \mathrm{d}t \quad (p > 2)$$

among all trajectories u with prescribed initial conditions, converge as $\varepsilon \to 0$ to a solution of the semilinear wave equation

$$u'' - \Delta u + |u|^{p-2}u = 0 \ \text{ in } \mathbb{R}^d \times [0, \infty[.$$

This conjecture has been checked positively for a finite time horizon $T < \infty$ first in [Ste11] and then for $T = \infty$ by Serra & Tilli in [SeT12]. Already in [De 96, Rem. 1] it is speculated that some similar result could hold for more general functionals of the Calculus of Variations as well. We proceed here by refining the analysis of [Ste11] in order to take into account dissipative effects $\nu > 0$ as well. The outcome of this extension is a theory which is indeed independent of the character of equation (5.1), provided $\rho + \nu > 0$. This is a quite remarkable feature of the WIDE formalism which in principle could make it of use in relation with a significant range of evolution problems. We exploit this fact in Subsection 6.6 where the limits $\rho \to 0$ and $\nu \to 0$ are discussed by means of a Γ-convergence analysis.

Indeed, the classical variational formulations of equations (5.1) for $\nu = 0$, that can be found in the literature (see for example [ShS98]), are based on the Euler-Lagrange equation of the action functional

$$\boldsymbol{H}[u] = \int_{t_\mathrm{a}}^{t_\mathrm{b}} \int_\Omega \left(\frac{\rho}{2}|u'|^2 - \frac{1}{2}|\nabla u|^2 - F(u) \right) \mathrm{d}x \, \mathrm{d}t.$$

The WIDE variational approach differs from this principle in some crucial points. First, Hamilton's principle is indeed a *stationarity principle* for it generally corresponds to the quest for a saddle point of the action functional (note however that this will be a true minimum for small t_b). On the contrary, the WIDE principle relies on a true constrained *minimization*. Moreover, the WIDE principle is not invariant by time reversal. This is indeed crucial as the WIDE perspective is naturally incorporating dissipative effects thus qualifying it as a suitable tool in order to discuss limiting mixed dissipative/nondissipative dynamics. Note that dissipative effects cannot be directly treated via the stationarity principle related to $\boldsymbol{H}$, and one resorts in considering the classical Lagrange-D'Alembert principle instead. Finally, in classical mechanics Hamilton's approach calls for the specification of an artificial finite-time interval $]t_\mathrm{a}, t_\mathrm{b}[$ and a final state. In particular, the WIDE functionals directly encode directionality of time by explicitly requiring the knowledge of just *initial states*. The price to pay within the WIDE functional method is the check of the extra limit $\varepsilon \to 0$. This is exactly the main object of the following chapters.

Let us mention here that other variational principles for characterizing entire trajectories of evolutionary systems are available; In the case of linear systems we refer to Biot's work on irreversible Thermodynamics [Bio55] and GURTIN's principle for viscoelasticity and elastodynamics [Gur63, Gur64a, Gur64b] among others (see also the survey by HLAVÁČEK [Hla69]). In the nonlinear case, a crucial result is the BRÉZIS, EKELAND, & NAYROLES principle [BrE76a, BrE76b, Nay76a, Nay76b] (see also [Rou05, Theorem 8.93], the monograph [Gho09], and the papers [Ste08a, Ste08b, Ste09]).

Review of the literature on weighted functionals

Global-in-time minimization of weighted functionals has been already considered in the purely dissipative (viscous) case. In particular, this functional approach has been developed for so called *Weighted Energy-Dissipation (WED) functionals*

$$u \mapsto \int_0^T \mathrm{e}^{-t/\varepsilon}\Big[\varepsilon\Psi(u') + \mathcal{E}(u)\Big]\,\mathrm{d}t$$

where Ψ is a suitable nonnegative and convex dissipation potential. In the linear case $\Psi(u') = \|u'\|^2/2$, some results can be found in the classical monograph by LIONS & MAGENES [LiM72]. As for the nonlinear case, this procedure has been followed by ILMANEN [Ilm94] for proving existence and partial regularity of the so-called Brakke mean curvature flow of varifolds.

Results and applications to rate-independent dissipative systems $\Psi(u') = \|u'\|$ have been presented by MIELKE & ORTIZ [MiO08] and then extended and coupled with time-discretization in [MiS08]. For λ-convex energies $\mathcal{E}$, the convergence proof $u_\varepsilon \to u$ in Hilbert and metric spaces has been provided in [MiS11] and [RS*11a, RS*11b], respectively. An application in the context of gradient flows driven by linear-growth functionals and, in particular, to mean curvature flow of graphs is given in [SpS11].

Two examples of relaxation of gradient flows related to microstructure evolution are provided by CONTI & ORTIZ [CoO08]. There the energy fails to be lower semi-continuous and thus also the WED functional. Nevertheless, it was shown that relaxations can be rigorously derived.

An application to crack propagation is given by LARSEN, ORTIZ, & RICHARDSON [LOR09]. Eventually, the doubly nonlinear case $\Psi(u') = \|u'\|^p/p$ ($p > 2$) is addressed in [AkS10, AkS11]. A duality-based WED approach to another large class of nonlinear evolutions including the two-phase Stefan problem and the porous-media equation is presented in [AkS12].

Eventually, a similar functional approach (with ε fixed though) has been considered by LUCIA, MURATOV, & NOVAGA in connection with traveling waves in reaction-diffusion-advection problems [LMN08, MuN08a, MuN08b].

As already pointed out above, the Weighted-Inertia-Energy approach is based on a conjecture of DE GIORGI [De 96], which was positively answered in [Ste11] and [SeT12] for the specific case $F(u) = |u|^p/p$ and without additional dissipative term.

6 The WIDE principle for a final time horizon

In this chapter we present the results of [LiS13b], where the WIDE functional $\boldsymbol{W}_\varepsilon$ in (5.2) was studied in the case of a finite time interval, i.e., for $T < \infty$. As a main result of this chapter we show that limits of minimizers of the WIDE functional $\boldsymbol{W}_\varepsilon$ are solutions of the limit equation (5.1). The proof of this result rests upon the validity of an a priori estimate on the minimizers of the WIDE functional and is obtained by considering a time-discrete version of the WIDE principle. This time-discrete version was already briefly discussed in the introduction and might be of independent interest.

In Section 6.2 we specify the assumptions on the ingredients, i.e., initial conditions, potential F, etc., and introduce the functional analytic setting. Moreover, we present the main result in Theorem 6.2.2. Next, we discuss the existence (and uniqueness) of minimizers of the WIDE functional in Section 6.3. Here, the existence and uniqueness of minimizer follow from the Direct Method of Calculus of Variations once we have shown that the WIDE functional is (uniformly) convex. For this we follow the ideas in [MiS11] and assume that the potential F is λ-convex on $\mathbb{R}$ (see (6.8)). In Section 6.4 we prove the main result: the convergence of minimizers of the WIDE functional to solutions of the limit equation (5.1). Here, we only give a formal derivation of the crucial estimate, the rigorous and rather technical derivation is postponed to the final Section 6.5. In particular, in Section 6.5 we introduce the time-discrete version of the WIDE principle whose well-posedness is addressed in Subsection 6.5.1. At the discrete level we are able to mimic the formal derivation of the key estimate of Section 6.4 in a rigorous fashion (see Subsection 6.5.2) and use the Γ-convergence of the discrete to the continuous WIDE functionals in order to pass the discrete estimate to the continuous case in Subsection 6.5.3.

The present variational formalism is well-suited in order to describe limiting behaviors. In particular, we comment in Section 6.6 on the possibility of considering from a variational viewpoint the limits $\rho \to 0$ and $\nu \to 0$. This will be done within the classical frame of Mosco convergence, resp. Γ-convergence. Finally, in Section 6.7 we consider the case of a finite-dimensional state space, where we are able to prove sharp estimates for the convergence of the minimizers.

6.1 Formal derivation of the variational principle

Following the presentation in [MiO08, Sect. 2] we shall provide a formal derivation of the WIDE principle.

A possible approach to solve problems of the form (5.1) for a finite reference time T

is time-discretization. Specifically, suppose that we are given the initial state u^0 and the initial velocity u^1 at time $t_0 = 0$ and wish to approximate the solution $U_n \approx u(t_n)$ at times $t_n = n\,\tau$, with $n = 2, \dots, N$ and $\tau = T/N$. Given $U_0 = u^0$ and $U_1 = u^0 + \tau u^1$ the time incremental version of (5.1) then reads: Find U_n such that

$$\rho \frac{U_n - 2U_{n-1} + U_{n-2}}{\tau^2} + \nu \frac{U_n - U_{n-1}}{\tau} + \mathrm{D}\mathcal{E}(U_n) = 0, \quad \text{for } n = 2, \dots, N.$$

Introducing the notation $\boldsymbol{\delta} U_n = (U_n - U_{n-1})/\tau$ and $\boldsymbol{\delta}^2 U_n = \boldsymbol{\delta}(\boldsymbol{\delta} U_n)$ for the first and second order difference quotient we see that the sequence of equations above is equivalent to the following sequence of minimization problems:

$$\begin{aligned} &U_n \in \operatorname*{Arg\,min}_V \boldsymbol{\Phi}_\tau(U_{n-2}, U_{n-1}, V), \quad n = 2, \dots, N, \qquad \text{where} \\ &\boldsymbol{\Phi}_\tau(U_{n-2}, U_{n-1}, U_n) = \frac{\rho}{2} \|\boldsymbol{\delta}^2 U_n\|^2 + \frac{\nu}{2\tau} \Big(\|\boldsymbol{\delta} U_n\|^2 - \|\boldsymbol{\delta} U_{n-1}\|^2 \Big) \\ &\qquad\qquad + \frac{1}{\tau^2} \Big(\mathcal{E}(U_n) - 2\mathcal{E}(U_{n-1}) + \mathcal{E}(U_{n-2}) \Big). \end{aligned} \tag{6.1}$$

The incremental functional $\boldsymbol{\Phi}_\tau$ combines energy and kinetics. We emphasize that the problems in (6.1) are solved *causally*: The problem for $n = 2$ is solved first with initial conditions U_0, U_1 in order to compute U_2. Subsequently, problem $n = 3$ is solved taking the solution U_2 of the preceding problem and U_1, and so on. We note that the additional terms $-\frac{\nu\tau}{2}\|\boldsymbol{\delta} U_{n-1}\|^2$ and $-2\mathcal{E}(U_{n-1}) + \mathcal{E}(U_{n-2})$ in the definition of $\boldsymbol{\Phi}_\tau$ are added such that the kinetic and energy terms are of the same order in τ.

Instead of solving each of the minimization problems separately we want to collect the incremental problems in (6.1) into a single minimum problem for the whole trajectory $\boldsymbol{U} = (U_0, \dots, U_N)$. In the theory of optimization the simultaneous optimization of two or more conflicting objectives (subject to certain constraints) is known as multi-criteria optimization (see e.g. [Cla90]).

Considering the minimum problems in (6.1) as problems in the entire trajectory $\boldsymbol{U}$, i.e., minimizing $\boldsymbol{U} \mapsto \boldsymbol{\Phi}_\tau^n[\boldsymbol{U}] = \boldsymbol{\Phi}_\tau(U_{n-2}, U_{n-1}, U_n)$ for $n = 2, \dots, N$, we see that the nth problem is coupled to the $(n \pm k)$th problem for $k = 1, 2$. To overcome this problem we combine all of the functionals $\boldsymbol{\Phi}_\tau^n$ into a single objective functional $\widetilde{\boldsymbol{W}}_\tau$, called *aggregate objective functional (AOF)*. In its simplest form it is given as the weighted sum of the functionals $\boldsymbol{\Phi}_\tau^n$

$$\widetilde{\boldsymbol{W}}_\tau[\boldsymbol{U}] = \sum_{n=2}^N \tau e_n \boldsymbol{\Phi}_\tau^n[\boldsymbol{U}] = \sum_{n=2}^N \tau e_n \boldsymbol{\Phi}_\tau(U_{n-2}, U_{n-1}, U_n),$$

where $\boldsymbol{e} = (e_2, \dots, e_N)$ are positive *Pareto weights*.

Obviously, since the nth step influences the $(n-1)$th, $(n-2)$th, and so on, we have lost causality. In order to ensure causality (at least in a relaxed sense) we choose the Pareto weights in such a way that the minimization of $\widetilde{\boldsymbol{W}}_\tau$ with respect to the entire trajectory $\boldsymbol{U}$ is (almost) equivalent to the incremental solution of (5.5). This is done by choosing the Pareto weights such that $e_2 \gg e_3 \gg \ldots$, which gives a much larger importance to the first incremental problem relative to the second, the second relative to the third and so on. Practically we achieve this by letting the positive weights depend on a parameter $\varepsilon > 0$ such that

$$\lim_{\varepsilon\to 0} \frac{e^\varepsilon_{n+1}}{e^\varepsilon_n} = 0. \tag{6.2}$$

Inserting these weights into $\widetilde{\boldsymbol{W}}_\tau[\boldsymbol{U}]$ gives

$$\widetilde{\boldsymbol{W}}_{\tau,\varepsilon}[\boldsymbol{U}] = \sum_{n=2}^{N} e^\varepsilon_n \tau \left[\frac{\rho}{2}\|\boldsymbol{\delta}^2 U_n\|^2 + \boldsymbol{\delta}\left(\frac{\nu}{2}(\|\boldsymbol{\delta} U_n\|^2)\right) + \boldsymbol{\delta}^2\mathcal{E}(U_n)\right].$$

Hence, assuming that the discrete weights $\boldsymbol{e}_\varepsilon$ satisfy $e^\varepsilon_n \approx e_\varepsilon(t_n)$, for a given positive and sufficiently smooth function $e_\varepsilon : [0,T] \to]0,\infty[$, we see that $\widetilde{\boldsymbol{W}}_{\tau,\varepsilon}$ is the time-discrete version of the functional given by

$$\widetilde{\boldsymbol{W}}_\varepsilon[u] = \int_0^T e_\varepsilon(t)\left[\frac{\rho}{2}\|u''(t)\|^2 + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\nu}{2}\|u'(t)\|^2\right) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{E}(u(t))\right]\mathrm{d}t.$$

Integration by parts gives the functional in the more familiar form

$$\widetilde{\boldsymbol{W}}_\varepsilon[u] = \int_0^T \left[e_\varepsilon(t)\frac{\rho}{2}\|u''(t)\|^2 - e'_\varepsilon(t)\frac{\nu}{2}\|u'(t)\|^2 + e''_\varepsilon(t)\mathcal{E}(u(t))\right]\mathrm{d}t + \boldsymbol{B}^{0,T}_\varepsilon[u],$$
$$\text{where}\quad \boldsymbol{B}^{0,T}_\varepsilon[u] = \left[\frac{e_\varepsilon\nu}{2}\|u'\|^2 + e_\varepsilon\langle \mathrm{D}\mathcal{E}(u), u'\rangle - e'_\varepsilon\mathcal{E}(u)\right]_0^T.$$

While causality requires that $t \mapsto e_\varepsilon(t)$ is monotonically decreasing, the limiting condition (6.2) means that $e_\varepsilon(b)/e_\varepsilon(a) \to 0$ for all $0 \le a < b \le T$ as ε tends to 0. Therefore, as we have that

$$\frac{1}{e_\varepsilon(a)}\frac{e_\varepsilon(b) - e_\varepsilon(a)}{b-a} = \frac{1}{b-a}\left(\frac{e_\varepsilon(b)}{e_\varepsilon(a)} - 1\right)$$

it holds that $e'_\varepsilon(t)/e_\varepsilon(t) \to -\infty$ for $\varepsilon \to 0$ and $t \in [0,T]$. Here, an admissible and particular simple choice is obtained by assuming that

$$\frac{e'_\varepsilon(t)}{e_\varepsilon(t)} = -\frac{1}{\varepsilon} \quad \text{which gives} \quad t \mapsto e_\varepsilon(t) = \mathrm{e}^{-t/\varepsilon},$$

where we have set $e_\varepsilon(0) = 1$ for definiteness. For this particular choice of the weight function the functional $\widetilde{\boldsymbol{W}}_\varepsilon$ reads

$$\widetilde{\boldsymbol{W}}_\varepsilon[u] = \int_0^T \mathrm{e}^{-t/\varepsilon}\left[\frac{\rho}{2}\|u''\|^2 + \frac{\nu}{2\varepsilon}\|u'\|^2 + \frac{1}{\varepsilon^2}\mathcal{E}(u)\right]\mathrm{d}t + \boldsymbol{B}^{0,T}_\varepsilon[u],$$
$$\text{where}\quad \boldsymbol{B}^{0,T}_\varepsilon[u] = \left[\mathrm{e}^{-t/\varepsilon}\Big(\langle \mathrm{D}\mathcal{E}(u), u'\rangle + \frac{1}{\varepsilon}\mathcal{E}(u) + \frac{\nu}{2}\|u'\|^2\Big)\right]_0^T.$$

Hence, we obtained $\widetilde{\boldsymbol{W}}_\varepsilon = \boldsymbol{W}_\varepsilon/\varepsilon^2 + \boldsymbol{B}_\varepsilon^{0,T}$, namely the WIDE functional $\boldsymbol{W}_\varepsilon$, introduced in (5.2), with an additional boundary contribution.

We immediately check that the Euler-Lagrange equation for $\widetilde{\boldsymbol{W}}_\varepsilon$ is indeed given by (5.3) while due to the additional $\boldsymbol{B}_\varepsilon^{0,T}$ we have the final conditions

$$\left.\begin{aligned} \rho u'' + \nu u' + \mathrm{D}\mathcal{E}(u) = 0 \\ -\varepsilon\rho u''' + \mathrm{D}^2\mathcal{E}(u)u' + \rho u'' + \nu u' + \mathrm{D}\mathcal{E}(u) = 0 \end{aligned}\right\} \quad \text{for } t = T. \tag{6.3}$$

Comparing these final conditions to (5.4) we see that the additional boundary term $\boldsymbol{B}_\varepsilon^{0,T}$ has an essential impact on the form of the final conditions. In particular, the first equation is enforcing, independently of ε, the attainment of the limit equation (5.1) at the final time T. The second final condition, which simplifies to $\mathrm{D}^2\mathcal{E}(u(T))u'(T) - \varepsilon\rho u'''(T) = 0$ does not seem to have a particular meaning.

For the rest of this work we will neglect the boundary term $\boldsymbol{B}_\varepsilon^{0,T}$ and consider the functional $\boldsymbol{W}_\varepsilon$ instead of $\widetilde{\boldsymbol{W}}_\varepsilon$. Firstly, this is motivated by the sake of simplicity, as already in the case of the functional $\boldsymbol{W}_\varepsilon$ computations involve various boundary terms which are hard to treat. Secondly, at least in the infinite dimensional case the quantity $\mathrm{D}^2\mathcal{E}(u)u'$ turns out to be not well-defined in the natural energy space.

We justify the negligence of $\boldsymbol{B}_\varepsilon^{0,T}$ by assuming that T/ε is sufficiently large such that the factor $\exp(-T/\varepsilon)$ is small. However, in [MiS11, Section 5.7] it was shown that in the parabolic case $\rho = 0$, and for a simple ODE example, sharper convergence estimates could be obtained when considering the functional $\widetilde{\boldsymbol{W}}_\varepsilon$, and hence $\boldsymbol{B}_\varepsilon^{0,T}$, instead of $\boldsymbol{W}_\varepsilon$.

6.2 Preliminaries and main result

We shall start by recalling some assumptions which will be tacitly assumed throughout the remainder of this chapter. Moreover, we introduce our weak solution notion for the Euler-Lagrange equation (5.3)–(5.4). In particular, let us consider $\Omega \subset \mathbb{R}^d$, $d \geq 1$, which is assumed to be an open and bounded Lipschitz domain. Concerning the potential $F \in \mathrm{C}^1(\mathbb{R})$ we ask for $f = F' : \mathbb{R} \to \mathbb{R}$ to be of polynomial growth. More precisely, we ask for some constant $C > 0$ and $p \geq 2$ such that for all $u \in \mathbb{R}$

$$\frac{1}{C}|u|^p \leq F(u) + C \quad \text{and} \quad |f(u)|^{p'} \leq C(1 + |u|^p), \tag{6.4}$$

where $1/p + 1/p' = 1$. Note that these growth assumptions imply that F has at most p-growth. Moreover, we assume that the stiffness matrix $\mathbb{A} \in \mathrm{L}^\infty(\Omega; \mathbb{R}^{d\times d})$ is symmetric and that there exists a constant $\gamma_\mathbb{A} > 0$ such that

$$\text{for all } \xi \in \mathbb{R}^d \text{ and for almost all } x \in \Omega: \qquad \xi \cdot \mathbb{A}(x)\xi \geq \gamma_\mathbb{A}|\xi|^2. \tag{6.5}$$

Hence, we define the Hilbert space $H = \mathrm{L}^2(\Omega)$ and the Banach spaces $X = \mathrm{L}^p(\Omega)$ and $Z = \mathrm{H}_0^1(\Omega)$. Obviously, we have $Z \subset H$ compactly and $X \subset H$ continuously. Let $\langle \cdot, \cdot \rangle$ denote the duality pairing both on $Z^* \times Z$ and $X^* \times X$ and the usual scalar product on

H. Moreover, $\|\cdot\|$ denotes the norm on H and $\|\cdot\|_B$ stands for the norm of a normed space B.

We define the (*energy*) functional $\mathcal{E} : Z \cap X \to \mathbb{R}$ as

$$\mathcal{E}(u) = \int_\Omega \left[\frac{1}{2}\nabla u \cdot \mathbb{A}\nabla u + F(u)\right] \mathrm{d}x$$

and the linear and nonlinear operators $\mathcal{A} : Z \to Z^*$ and $\mathcal{B} : X \to X^*$ as

$$\langle \mathcal{A}u, v\rangle = \int_\Omega \nabla u{\cdot}\mathbb{A}\nabla v \; \mathrm{d}x, \quad \text{and} \quad \langle \mathcal{B}(u), v\rangle = \int_\Omega f(u)v \; \mathrm{d}x$$

such that $\langle \mathcal{A}u, u\rangle \geq \gamma_{\mathbb{A}}\|u\|_Z^2$ and $\mathcal{B}(u) = f(u)$ almost everywhere. Moreover, using the growth conditions for f we have that $\mathrm{D}\mathcal{E} = \mathcal{A}{+}\mathcal{B} : Z \cap X \to Z^*{+}X^*$ being bounded.

We introduce the evolution spaces

$$\begin{aligned} &\mathbb{V} = \mathrm{H}^1(0,T;H) \cap \mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X), \\ \text{and} \quad &\mathbb{Y} = \Big\{u \in \mathbb{V} \;\mathbf{:}\; \rho u' \in \mathrm{H}^1(0,T;H)\Big\}. \end{aligned} \tag{6.6}$$

Here we used the notation $\rho u'$ in the definition of the space $\mathbb{Y}$ to highlight that we also consider the parabolic case $\rho = 0$ in which the space $\mathbb{Y}$ and $\mathbb{V}$ coincide.

Finally, we assume that the operator $\mathcal{B}$ is weakly continuous on $\mathbb{V}$, i.e., we have that

$$u_k \rightharpoonup u \quad \text{in} \quad \mathbb{V} \quad \Longrightarrow \quad \mathcal{B}(u_k) \rightharpoonup \mathcal{B}(u) \quad \text{in} \quad \mathbb{V}^*. \tag{6.7}$$

A choice for the function $f : \mathbb{R} \to \mathbb{R}$ fulfilling the growth assumptions in (6.4) and the weak continuity property in (6.7) is $f(u) = |u|^{p-2}u + \ell(u)$ where $\ell \in \mathrm{C}^{0,1}(\mathbb{R})$.

Additionally, let us assume that F is λ-convex (as a function on $\mathbb{R}$) for some given $\lambda \in \mathbb{R}$, i.e.,

$$u \mapsto F(u) - \frac{\lambda}{2}|u|^2 \quad \text{is convex.} \tag{6.8}$$

Equivalently, F is λ-convex if and only if

$$F(u_\theta) \leq \theta F(u_1) + (1{-}\theta)F(u_0) - \frac{\lambda\theta(1{-}\theta)}{2}|u_0{-}u_1|^2$$

for all $u_0, u_1 \in \mathbb{R}$, $u_\theta = \theta u_1{+}(1{-}\theta)u_0$ and $\theta \in [0,1]$. Moreover, we have the convexity estimate

$$\forall\, u, v \in \mathbb{R} : \qquad F(v) \geq F(u) + F'(u)\cdot(v{-}u) + \frac{\lambda}{2}|v{-}u|^2. \tag{6.9}$$

We will see in the following section, that the λ-convexity of F yields the uniform convexity of the WIDE functional and thus the existence of (unique) minimizers.

Before stating our main result of this chapter let us recall the definition of the WIDE functional $\boldsymbol{W}_\varepsilon : \mathbb{Y} \to \mathbb{R}$ and the nonempty, closed, and convex set $\mathbb{K}(u^0, u^1)$ which encodes the initial conditions

$$\boldsymbol{W}_\varepsilon[u] = \int_0^T \mathrm{e}^{-t/\varepsilon} \left[\frac{\varepsilon^2\rho}{2}\|u''\|^2 + \frac{\varepsilon\nu}{2}\|u'\|^2 + \mathcal{E}(u)\right] \mathrm{d}t,$$

$$\text{and} \quad \mathbb{K}(u^0,u^1) = \left\{u \in \mathbb{Y} \,:\, u(0)=u^0,\ \rho u'(0) = \rho u^1\right\}.$$

Our analysis, in particular the derivation of a priori estimates, relies on the specific structure of the Euler–Lagrange equation for $\boldsymbol{W}_\varepsilon$. Let u_ε minimize $\boldsymbol{W}_\varepsilon$ in $\mathbb{K}(u^0,u^1)$. By considering $h \mapsto \boldsymbol{W}_\varepsilon[u_\varepsilon{+}hv]$ for $v \in \mathbb{K}(0,0)$ we obtain that

$$\forall\, v \in \mathbb{K}(0,0): \quad \int_0^T \mathrm{e}^{-t/\varepsilon}\Big[\varepsilon^2\rho\langle u_\varepsilon'', v''\rangle + \varepsilon\nu\langle u_\varepsilon', v'\rangle + \langle \mathrm{D}\mathcal{E}(u_\varepsilon), v\rangle\Big]\,\mathrm{d}t = 0. \tag{6.10}$$

Hence, we have the following.

Lemma 6.2.1 (Euler-Lagrange equation) *Let u_ε be the unique minimizer of the functional $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}(u^0,u^1)$. Then, u_ε (formally) solves*

$$\varepsilon^2\rho u_\varepsilon'''' - 2\varepsilon\rho u_\varepsilon''' + (\rho{-}\varepsilon\nu)u_\varepsilon'' + \nu u_\varepsilon' + \mathrm{D}\mathcal{E}(u_\varepsilon) = 0 \quad \textit{for almost all } t \in \,]0,T[\,, \tag{6.11a}$$

subjected to the initial and final conditions

$$u_\varepsilon(0) = u^0, \quad \rho u_\varepsilon'(0) = \rho u^1, \quad \textit{and} \tag{6.11b}$$

$$\varepsilon^2\rho u_\varepsilon''(T) = 0, \quad \varepsilon^2\rho u_\varepsilon'''(T) = \varepsilon\nu u_\varepsilon'(T). \tag{6.11c}$$

Now, we can formulate the main theorem of this chapter, whose proof is postponed to Section 6.4.2.

Theorem 6.2.2 (WIDE principle) *Let u_ε be a minimizer of the* WIDE *functional $\boldsymbol{W}_\varepsilon$ on the nonempty, closed, and convex set $\mathbb{K}(u^0,u^1)$, then, there exists a (not relabeled) subsequence u_ε such that $u_\varepsilon \rightharpoonup u$ in $\mathbb{V}$, where u solves*

$$\rho u'' + \nu u' + \mathrm{D}\mathcal{E}(u) = 0, \qquad u(0) = u^0,\ \rho u'(0) = \rho u^1. \tag{6.12}$$

6.3 Well-posedness of the minimum problem

In this section we show the existence of minimizers u_ε of the WIDE functional $\boldsymbol{W}_\varepsilon$. More precisely, we show that for ε sufficiently small $\boldsymbol{W}_\varepsilon$ turns out to be uniformly convex in $\mathrm{H}^1(0,T;H) \cap \mathrm{L}^2(0,T;Z)$ (resp. $\mathrm{H}^2(0,T;H)$ for $\rho > 0$). In the convex case, i.e., $\lambda^- = \max\{0,-\lambda\} = 0$, the existence of a (unique) minimizer is a direct consequence of the Direct Method. As for the general nonconvex case $\lambda^- > 0$, existence and uniqueness of minimizers follow by letting ε be small enough.

Theorem 6.3.1 (Well-posedness of minimum problem) *For small $\varepsilon > 0$ the functional $\boldsymbol{W}_\varepsilon$ is uniformly convex on $\mathbb{K}(u^0,u^1)$ with respect to the metric of* $\mathrm{H}^1(0,T;H) \cap$

$\mathrm{L}^2(0,T;Z)$ *(resp.* $\mathrm{H}^2(0,T;H)$ *for* $\rho>0$*), i.e., there exists* $\kappa_\varepsilon>0$ *such that*

$$\boldsymbol{W}_\varepsilon[u_\theta] \le (1-\theta)\boldsymbol{W}_\varepsilon[u_0] + \theta\boldsymbol{W}_\varepsilon[u_1] - \frac{\kappa_\varepsilon\theta(1-\theta)}{2}\Big(\|u_0-u_1\|^2_{\mathrm{H}^k(0,T;H)} + \|u_0-u_1\|^2_{\mathrm{L}^2(0,T;Z)}\Big)$$

for $u_0,u_1\in\mathbb{K}(u^0,u^1)$*,* $u_\theta=(1-\theta)u_0+\theta u_1$*,* $\theta\in[0,1]$ *and* $k=2$ *for* $\rho>0$ *and* $k=1$ *otherwise. In particular,* $\boldsymbol{W}_\varepsilon$ *admits a unique minimizer* $u_\varepsilon\in\mathbb{K}(u^0,u^1)$*.*

Proof: As already mentioned the convex case $\lambda^-=0$ is quite straightforward so let us assume from the very beginning that $\lambda^->0$ and decompose $\boldsymbol{W}_\varepsilon$ into the sum of a quadratic part $\boldsymbol{Q}_\varepsilon$ and a convex remainder $\boldsymbol{R}_\varepsilon$ as follows

$$\begin{aligned}\boldsymbol{W}_\varepsilon[u] &= \int_0^T \frac{\mathrm{e}^{-t/\varepsilon}}{2}\Big\{\varepsilon^2\rho\|u''\|^2+\varepsilon\nu\|u'\|^2-\lambda^-\|u\|^2+\langle\mathcal{A}u,u\rangle\Big\}\,\mathrm{d}t + \int_0^T \mathrm{e}^{-t/\varepsilon}\mathcal{G}(u)\,\mathrm{d}t\\ &= \boldsymbol{Q}_\varepsilon[u]+\boldsymbol{R}_\varepsilon[u]\end{aligned}$$

with $\mathcal{G}(u)=\int_\Omega[F(u)-\lambda|u|^2/2]\,\mathrm{d}x$ convex by (6.8). In order to handle the quadratic part $\boldsymbol{Q}_\varepsilon$ we proceed as in [MiS11, Proof of Proposition 2.1] and exploit the auxiliary function $t\mapsto v(t)=\mathrm{e}^{-t/(2\varepsilon)}u(t)$ and readily check that

$$\mathrm{e}^{-t/(2\varepsilon)}u' = v'+\frac{1}{2\varepsilon}v, \quad\text{and}\quad \mathrm{e}^{-t/(2\varepsilon)}u'' = v''+\frac{1}{\varepsilon}v'+\frac{1}{4\varepsilon^2}v. \tag{6.13}$$

Note that, by possibly letting ε be small, standard computations ensure that we have the following estimates for the norms of u and v

$$\mathrm{e}^{-T/\varepsilon}\|u\|^2_{\mathrm{L}^2(0,T;Z)} \le \|v\|^2_{\mathrm{L}^2(0,T;Z)} \le \|u\|^2_{\mathrm{L}^2(0,T;Z)}, \tag{6.14a}$$

$$\varepsilon^4\mathrm{e}^{-T/\varepsilon}\|u\|^2_{\mathrm{H}^2(0,T;H)} \le \|v\|^2_{\mathrm{H}^2(0,T;H)} \le \varepsilon^{-4}\|u\|^2_{\mathrm{H}^2(0,T;H)}. \tag{6.14b}$$

Inserting (6.13) into $\boldsymbol{Q}_\varepsilon[u]$, we rewrite the latter in terms of v as

$$\begin{aligned}\boldsymbol{Q}_\varepsilon[u] &= \int_0^T\left(\frac{\rho\varepsilon^2}{2}\|v''\|^2+\frac{\rho+\varepsilon\nu}{2}\|v'\|^2+\frac{\rho+4\varepsilon\nu-16\varepsilon^2\lambda^-}{32\varepsilon^2}\|v\|^2+\frac{1}{2}\langle\mathcal{A}v,v\rangle\right)\mathrm{d}t\\ &\quad+\int_0^T\left(\rho\varepsilon\langle v'',v'\rangle+\frac{\rho}{4}\langle v'',v\rangle+\frac{\rho+2\varepsilon\nu}{4\varepsilon}\langle v',v\rangle\right)\mathrm{d}t.\end{aligned}$$

Using integration by parts in the mixed terms in the second integral we arrive at

$$\begin{aligned}\boldsymbol{Q}_\varepsilon[u] &= \int_0^T\left(\frac{\rho\varepsilon^2}{2}\|v''\|^2+\frac{\rho+2\varepsilon\nu}{4}\|v'\|^2+\frac{\rho+4\varepsilon\nu-16\varepsilon^2\lambda^-}{32\varepsilon^2}\|v\|^2+\frac{1}{2}\langle\mathcal{A}v,v\rangle\right)\mathrm{d}t\\ &\quad+\frac{\rho\varepsilon}{2}(\|v'(T)\|^2-\|v'(0)\|^2)+\frac{\rho}{4}\Big(\langle v'(T),v(T)\rangle-\langle v'(0),v(0)\rangle\Big)\\ &\quad+\frac{\rho+2\varepsilon\nu}{8\varepsilon}(\|v(T)\|^2-\|v(0)\|^2) = \boldsymbol{I}_\varepsilon[v]+\boldsymbol{B}_\varepsilon[v],\end{aligned} \tag{6.15}$$

where $\boldsymbol{I}_\varepsilon$ is the integral contribution and $\boldsymbol{B}_\varepsilon$ collects all boundary terms. Looking only at $\boldsymbol{I}_\varepsilon$ we see that by letting ε sufficiently small, namely

$$\varepsilon < \frac{1}{4}\max\left\{\sqrt{\frac{\rho}{\lambda^-}}, \frac{\nu}{\lambda^-}\right\}, \tag{6.16}$$

the quadratic form $\boldsymbol{I}_\varepsilon$ is convex. The same holds also for the functional $\boldsymbol{B}_\varepsilon$ for it is quadratic in $v'(T)$ and $v(t)$.

Let now $\theta \in [0,1]$ and $u_0, u_1 \in \mathbb{K}(u^0, u^1)$ be given. Moreover, for $i = 0, 1$ we define $v_i(t) = \mathrm{e}^{-t/(2\varepsilon)}u_i(t)$ and $w = v_0 - v_1$. For all ε small enough one deduces the existence of $\widetilde{\kappa}_\varepsilon > 0$ such that

$$\begin{aligned}\boldsymbol{Q}_\varepsilon[u_\theta] &\le \theta(\boldsymbol{I}_\varepsilon[v_1] + \boldsymbol{B}_\varepsilon[v_1]) + (1-\theta)(\boldsymbol{I}_\varepsilon[v_0] + \boldsymbol{B}_\varepsilon[v_0]) \\ &\quad - \frac{\widetilde{\kappa}_\varepsilon\theta(1-\theta)}{2}\left\{\rho\|w''\|^2_{\mathrm{L}^2(0,T;H)} + \|w\|^2_{\mathrm{H}^1(0,T;H)} + \|w\|^2_{\mathrm{L}^2(0,T;Z)}\right\}.\end{aligned}$$

By exploiting the first estimates in (6.14a) and (6.14b), we have proved that $\boldsymbol{Q}_\varepsilon$ is uniformly convex with respect to the metric of $\mathrm{H}^1(0,T;H) \cap \mathrm{L}^2(0,T;Z)$ (or even $\mathrm{H}^2(0,T;H)$ if $\rho > 0$). As $\boldsymbol{W}_\varepsilon = \boldsymbol{Q}_\varepsilon + \boldsymbol{R}_\varepsilon$ and $\boldsymbol{R}_\varepsilon$ is convex, the uniform convexity of $\boldsymbol{W}_\varepsilon$ and the existence of a unique minimizer $u_\varepsilon \in \mathbb{K}(u^0, u^1)$ ensue. □

6.4 A priori estimate and limit passage

The key step in the proof of Theorem 6.2.2 is to establish an integral energy estimate on the family of minimizers u_ε which is independent of ε. Henceforth, the symbol C stands for any constant depending on data and independent of ρ, ν, and ε (and, later, the time step τ) and possibly changing from line to line. We shall prove the following lemma.

Lemma 6.4.1 (A priori estimate) *Let u_ε be a minimizer of the* WIDE *functional $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}(u^0, u^1)$. Then, for all sufficiently small ε*

$$(\rho+\nu)\int_0^T \|u_\varepsilon'\|^2\,\mathrm{d}t + \int_0^T \mathcal{E}(u_\varepsilon)\,\mathrm{d}t \le C. \tag{6.17}$$

Note that, owing to the growth conditions (6.4), the latter estimate entails in particular that minimizers of $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}(u^0, u^1)$ are uniformly bounded in $\mathbb{V}$. This provides the necessary compactness in order to prove our main result Theorem 6.2.2. The rigorous proof of Lemma 6.4.1 is postponed to Section 6.5.3.

6.4.1 A formal argument

The proof of Lemma 6.4.1 will be achieved using a quite technical time-discretization scheme in Section 6.5. Let us however provide here a formal argument by assuming smoothness of the solutions u_ε of the Euler-Lagrange equation (6.11a) and focusing on the (more difficult) case $\rho > 0$ only.

Let v denote the function $t \mapsto v(t) = (1{+}T{-}t)(u_\varepsilon'(t){-}u^1)$. Although v is not an admissible test function in $\mathbb{K}(0,0)$ we nevertheless test (6.11a) by v and take the integral on $[0,T]$. We shall imitate this procedure at the time-discrete level in a rigorous fashion. By recalling the formula

$$\forall g \in \mathrm{L}^1(0,T), \quad \int_0^T (1{+}T{-}t)g(t)\,\mathrm{d}t = \int_0^T g(t)\,\mathrm{d}t + \int_0^T \Big(\int_0^t g(s)\,\mathrm{d}s\Big)\,\mathrm{d}t$$

we easily compute that for the fourth order term in (6.11a) we have that

$$\begin{aligned}\varepsilon^2\rho\int_0^T \langle u_\varepsilon'''', v\rangle\,\mathrm{d}t &= \varepsilon^2\rho\int_0^T \langle u_\varepsilon'''', u_\varepsilon'{-}u^1\rangle\,\mathrm{d}t + \varepsilon^2\rho\int_0^T\int_0^t \langle u_\varepsilon'''', u_\varepsilon'{-}u^1\rangle\,\mathrm{d}s\,\mathrm{d}t\\ &= \frac{(1{+}T)\varepsilon^2\rho}{2}\|u_\varepsilon''(0)\|^2 + \varepsilon^2\rho\langle u_\varepsilon'''(T), u_\varepsilon'(T){-}u^1\rangle - \frac{3\varepsilon^2\rho}{2}\int_0^T \|u_\varepsilon''\|^2\,\mathrm{d}t,\end{aligned}$$

where we have used integration by parts several times as well as the final and initial conditions $u_\varepsilon''(T) = 0$ and $u_\varepsilon'(0) = u^1$, respectively. Analogously, we use integration by parts for the third order term giving

$$\begin{aligned}-2\varepsilon\rho\int_0^T \langle u_\varepsilon''', v\rangle\,\mathrm{d}t &= -2\varepsilon\rho\int_0^T \langle u_\varepsilon''', u_\varepsilon'{-}u^1\rangle\,\mathrm{d}t - 2\varepsilon\rho\int_0^T\int_0^t \langle u_\varepsilon''', u_\varepsilon'{-}u^1\rangle\,\mathrm{d}s\,\mathrm{d}t\\ &= 2\varepsilon\rho\int_0^T \|u_\varepsilon''\|^2\,\mathrm{d}t + \varepsilon\rho\int_0^T\int_0^t \|u_\varepsilon''\|^2\,\mathrm{d}s\,\mathrm{d}t - \varepsilon\rho\|u_\varepsilon'(T){-}u^1\|^2,\end{aligned}$$

where we used the final condition $u_\varepsilon''(T) = 0$ again. Next, for the second order term we have

$$(\rho{-}\varepsilon\nu)\int_0^T \langle u_\varepsilon'', v\rangle\,\mathrm{d}t = \frac{\rho{-}\varepsilon\nu}{2}\|u_\varepsilon'{-}u^1\|^2 + \frac{\rho{-}\varepsilon\nu}{2}\int_0^T \|u_\varepsilon'{-}u^1\|^2\,\mathrm{d}t.$$

Moreover, using the chain rule for $t \mapsto \mathcal{E}(u_\varepsilon(t))$ we obtain

$$\begin{aligned}\int_0^T \langle \mathrm{D}\mathcal{E}(u_\varepsilon), v\rangle\,\mathrm{d}t &= \mathcal{E}(u_\varepsilon(T)) - (1{+}T)\mathcal{E}(u^0) + \int_0^T \mathcal{E}(u_\varepsilon)\,\mathrm{d}t\\ &\quad - \int_0^T \langle \mathrm{D}\mathcal{E}(u_\varepsilon), u^1\rangle\,\mathrm{d}t - \int_0^T\int_0^t \langle \mathrm{D}\mathcal{E}(u_\varepsilon), u^1\rangle\,\mathrm{d}s\ \mathrm{d}t.\end{aligned}$$

Finally, we sum up the equations and exploit the final conditions $u_\varepsilon''(T) = 0$ and $\varepsilon\rho u_\varepsilon'''(T) = \nu u_\varepsilon'(T)$ and arrive, for sufficiently small ε at the following estimate

$$\int_0^T \Big\{\nu\|u_\varepsilon'\|^2 + \rho\|u_\varepsilon'\|^2 + \varepsilon\|u_\varepsilon''\|^2 + \mathcal{E}(u_\varepsilon)\Big\}\,\mathrm{d}t \le C\left(1 + \int_0^T |\langle \mathrm{D}\mathcal{E}(u_\varepsilon), u^1\rangle|\,\mathrm{d}s\right). \quad (6.18)$$

Hence, by using the growth conditions (6.4) and Young's inequality we absorb the remaining terms on the right-hand side and have thus shown that for small ε estimate (6.17) holds.

6.4.2 Proof of the main result

Let us now come to the proof of the main result in Theorem (6.2.2). Let u_ε be the family of minimizers of the WIDE functionals $\boldsymbol{W}_\varepsilon$. Owing to Lemma 6.4.1 we can extract a (not relabeled) subsequence u_ε such that $u_\varepsilon \rightharpoonup u$ in $\mathbb{V}$, i.e.,

$$u_\varepsilon \rightharpoonup u \quad \text{in } \mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X) \cap \mathrm{H}^1(0,T;H).$$

In order to check that the limit u solves the limit equation in (6.12) we consider an arbitrary $w \in \mathrm{C}_c^\infty([0,T[;Z \cap X)$ (in particular $w(T) = w'(T) = w''(T) = 0$) and define the test function $t \mapsto v_\varepsilon(t) = \mathrm{e}^{t/\varepsilon} w(t) - h_\varepsilon(t)$, where $h_\varepsilon(t) = w(0) + t(w'(0) + \frac{1}{\varepsilon} w(0))$. In particular, $v_\varepsilon \in \mathbb{K}(0,0)$ holds and we have that

$$\begin{aligned} v_\varepsilon'(t) &= \mathrm{e}^{t/\varepsilon} w'(t) + \frac{1}{\varepsilon}\mathrm{e}^{t/\varepsilon} w(t) - h_\varepsilon'(t) \quad \text{and} \\ v_\varepsilon''(t) &= \mathrm{e}^{t/\varepsilon} w''(t) + \frac{2}{\varepsilon}\mathrm{e}^{t/\varepsilon} w'(t) + \frac{1}{\varepsilon^2}\mathrm{e}^{t/\varepsilon} w(t). \end{aligned}$$

Since $v_\varepsilon \in \mathrm{C}_c^\infty([0,T[;Z \cap X)$ is an admissible test function from the weak Euler–Lagrange equation (6.10) one obtains

$$\begin{aligned} 0 = &\int_0^T \Big\{ \langle \rho u_\varepsilon'', \varepsilon^2 w'' + 2\varepsilon w' \rangle + \varepsilon\nu \langle u_\varepsilon', w' - \mathrm{e}^{-t/\varepsilon} h_\varepsilon' \rangle - \langle \mathrm{D}\mathcal{E}(u_\varepsilon), \mathrm{e}^{-t/\varepsilon} h_\varepsilon \rangle \Big\} \,\mathrm{d}t \\ &+ \int_0^T \langle \rho u_\varepsilon'' + \nu u_\varepsilon' + \mathrm{D}\mathcal{E}(u_\varepsilon), w \rangle \,\mathrm{d}t. \end{aligned}$$

Hence, integrating by parts and reordering the terms we arrive at the equation

$$\begin{aligned} \int_0^T \Big\{ \langle \nu u_\varepsilon' + \mathrm{D}\mathcal{E}(u_\varepsilon), w \rangle - \rho \langle u_\varepsilon', w' \rangle \Big\} \,\mathrm{d}t &= \int_0^T \langle u_\varepsilon', \varepsilon^2 \rho w''' + 2\varepsilon \rho w'' - \varepsilon \nu w' \rangle \\ &+ \int_0^T \varepsilon\nu \langle u_\varepsilon', \mathrm{e}^{-t/\varepsilon} h_\varepsilon' \rangle + \langle \mathrm{D}\mathcal{E}(u_\varepsilon), \mathrm{e}^{-t/\varepsilon} h_\varepsilon \rangle \,\mathrm{d}t - \rho \langle u^1, \varepsilon^2 w''(0) + 2\varepsilon w'(0) + w(0) \rangle, \end{aligned}$$

where we have used the initial condition $u_\varepsilon'(0) = u^1$ and that w and its derivatives vanish at $t = T$. By passing to the limit as ε tends to 0 and using the weak continuity of the operator $\mathcal{B}$ in (6.7) we obtain that the limit u satisfies

$$\int_0^T \Big(\langle \nu u' + \mathrm{D}\mathcal{E}(u), w \rangle - \rho \langle u', w' \rangle \Big) \,\mathrm{d}t + \rho \langle u^1, w(0) \rangle = 0.$$

Namely, u solves the limit equation in (6.12), where u'' makes sense in the dual space $\mathrm{L}^2(0,T;Z^*) + \mathrm{L}^{p'}(0,T;X^*)$. The initial condition $u(0) = u^0$ follows from the precompactness of $(u_\varepsilon)_{\varepsilon>0}$ in $\mathbb{V}$, whereas the second initial condition $u'(0) = u^1$ follows by the weak formulation of the limit equation. $\square$

6.5 The time-discrete WIDE principle

The proof of Theorem 6.2.2 in the previous section rests upon the possibility of proving the key estimate in Lemma 6.4.1. In this section we rigorously derive the key estimate by investigating a time-discrete version of the WIDE principle. More precisely, we replace the functional $\boldsymbol{W}_\varepsilon$ by a time-discrete version $\boldsymbol{W}_{\varepsilon\tau}$, where $\tau = T/N$, for $N \in \mathbb{N}$, is the constant time step. By mimicking the formal approach of Section 6.4.1 we derive a corresponding estimate at the discrete level which is uniform in ε *and* τ. Finally, using the Γ-convergence of the discrete WIDE functionals $\boldsymbol{W}_{\varepsilon\tau}$ to $\boldsymbol{W}_\varepsilon$ in the weak topology of $\mathbb{V}$ when τ goes to zero we pass the discrete estimate to the continuous case (see Proposition 6.5.6).

From here on we directly focus on the situation $\rho > 0$, the case $\rho = 0$ being covered in [MiS11]. We start by introducing the space of discrete trajectories

$$\mathbb{Y}_\tau = \big\{(U_0, \dots, U_N) \in H^{N+1} \,:\, (U_2, \dots, U_{N-2}) \in (Z \cap X)^{N-3}\big\}.$$

Moreover, similar to Section 6.1 we define the discrete WIDE functional $\boldsymbol{W}_{\varepsilon\tau} : \mathbb{Y}_\tau \to \mathbb{R}$ on the discrete trajectories $\boldsymbol{U} = (U_0, \dots, U_N)$ by

$$\boldsymbol{W}_{\varepsilon\tau}[\boldsymbol{U}] = \frac{\varepsilon^2\rho}{2}\sum_{j=2}^{N} \tau e_\varepsilon^j \big\|\boldsymbol{\delta}^2 U_j\big\|^2 + \frac{\varepsilon\nu}{2}\sum_{j=2}^{N-1} \tau e_\varepsilon^{j+1}\big\|\boldsymbol{\delta} U_j\big\|^2 + \sum_{j=2}^{N-2} \tau e_\varepsilon^{j+2}\mathcal{E}(U_j).$$

Given a vector $(V_0, \dots, V_N)$ we use the notation $\boldsymbol{\delta} V_j = (V_j - V_{j-1})/\tau$ for its *discrete derivative*, and $\boldsymbol{\delta}^2 V_j = \boldsymbol{\delta}(\boldsymbol{\delta} V_j)$ and $\boldsymbol{\delta}^3 V_j = \boldsymbol{\delta}(\boldsymbol{\delta}^2 V_j)$ for the second and third order discrete derivative, respectively, and so on. Moreover, we introduce the *discrete weights* $\boldsymbol{e}_\varepsilon = (e_\varepsilon^0, \dots, e_\varepsilon^N)$ given by

$$e_\varepsilon^i = \left(\frac{\varepsilon}{\varepsilon + \tau}\right)^i \quad \text{for } i = 0, \dots, N. \tag{6.19}$$

These weights are nothing but the discrete version of the exponentially decaying weight $t \mapsto \mathrm{e}^{-t/\varepsilon}$ for which we have that $\boldsymbol{\delta} e_\varepsilon^i + e_\varepsilon^i/\varepsilon = 0$. Namely, $\boldsymbol{e}_\varepsilon$ is the solution of the constant time step implicit Euler discretization of the problem $e_\varepsilon' + e_\varepsilon/\varepsilon = 0$, with the initial condition $e_\varepsilon(0) = 1$.

Finally, the discrete counterpart of the convex set $\mathbb{K}(u^0, u^1)$ is denoted by $\mathbb{K}_\tau(u^0, u^1)$ and defined via

$$\mathbb{K}_\tau(u^0, u^1) = \Big\{\boldsymbol{U} \in \mathbb{Y}_\tau \,:\, U_0 = u^0,\ \rho\boldsymbol{\delta} U_1 = \rho u^1\Big\}.$$

The unique minimizer of the time-discrete functional $\boldsymbol{W}_{\varepsilon\tau}$ on $\mathbb{K}_\tau(u^0, u^1)$ solves the corresponding Euler–Lagrange system and we directly compute that the latter reads

$$0 = \varepsilon^2\rho\sum_{j=2}^{N} \tau e_\varepsilon^j \langle\boldsymbol{\delta}^2 U_j, \boldsymbol{\delta}^2 V_j\rangle + \varepsilon\nu\sum_{j=2}^{N-1} \tau e_\varepsilon^{j+1}\langle\boldsymbol{\delta} U_j, \boldsymbol{\delta} V_j\rangle + \sum_{j=2}^{N-2} \tau e_\varepsilon^{j+2}\langle \mathrm{D}\mathcal{E}(U_j), V_j\rangle \tag{6.20}$$

for all $\boldsymbol{V} \in \mathbb{K}_\tau(0, 0)$. Analogously to the continuous case we use summation by parts to obtain the following result. (Here we omit the lengthy computations, an interested reader is referred to [LiS13b])

Lemma 6.5.1 (Discrete Euler-Lagrange equation) *Let $\boldsymbol{U}_{\varepsilon\tau} = (U_0^{\varepsilon\tau}, \dots, U_N^{\varepsilon\tau})$ be a minimizer of the discrete* WIDE *functional $\boldsymbol{W}_{\varepsilon\tau}$ on $\mathbb{K}_\tau(u^0, u^1)$. Then, $\boldsymbol{U}_{\varepsilon\tau}$ solves*

$$\varepsilon^2 \rho\, \boldsymbol{\delta}^4 U_{j+2}^{\varepsilon\tau} - 2\varepsilon\rho\, \boldsymbol{\delta}^3 U_{j+1}^{\varepsilon\tau} + \rho\, \boldsymbol{\delta}^2 U_j^{\varepsilon\tau} - \varepsilon\nu\, \boldsymbol{\delta}^2 U_{j+1}^{\varepsilon\tau} + \nu\, \boldsymbol{\delta} U_j^{\varepsilon\tau} + \mathrm{D}\mathcal{E}(U_j^{\varepsilon\tau}) = 0, \qquad (6.21a)$$

for $j = 2, \dots, N-2$ and subject to the initial and final conditions

$$U_0^{\varepsilon\tau} = u^0, \quad \textit{and} \quad \rho\boldsymbol{\delta} U_1^{\varepsilon\tau} = \rho u^1, \qquad (6.21b)$$

$$\varepsilon^2\rho\boldsymbol{\delta}^2 U_N^{\varepsilon\tau} = 0, \quad \textit{and} \quad \varepsilon\rho\boldsymbol{\delta}^2 U_{N-1}^{\varepsilon\tau} + \varepsilon\nu\boldsymbol{\delta} U_{N-1}^{\varepsilon\tau} = \varepsilon^2\rho\boldsymbol{\delta}^3 U_N^{\varepsilon\tau}. \qquad (6.21c)$$

Obviously, equations (6.21a)–(6.21c) are the discrete analog of equations (6.11a)–(6.11c).

6.5.1 Well-posedness of the discrete minimum problem

Exactly as in the time-continuous situation, in case that F is λ-convex, the functional $\boldsymbol{W}_{\varepsilon\tau}$ turns out to be uniformly convex for sufficiently small ε. Note that for all discrete trajectories $\boldsymbol{U} \in \mathbb{K}_\tau(u^0, u^1)$ we easily obtain the discrete Poincaré-type estimate

$$\sum_{k=2}^{N} \tau \|U_k\|^2 \le C \left(\|u^0\|^2 + \|u^1\|^2 + \sum_{k=2}^{N} \tau \|\boldsymbol{\delta}^2 U_k\|^2 \right), \qquad (6.22)$$

where C depends on T. Hence, the functional $\boldsymbol{W}_{\varepsilon\tau}$ is coercive on $\mathbb{K}_\tau(u^0, u^1)$. Indeed, the coercivity of $\boldsymbol{W}_{\varepsilon\tau}$ in $(Z \cap X)^{N-3}$ with respect to $(U_2, \dots, U_{N-2})$ is immediate. As for the coercivity in H we see that, due to (6.22), the discrete WIDE functional $\boldsymbol{W}_{\varepsilon\tau}$ controls the norm in H (up to constants depending on $T, \rho, \nu, \varepsilon$, and τ).

Remark 6.5.2 *Although the discrete* WIDE *functional $\boldsymbol{W}_{\varepsilon\tau}$ is only coercive on the set $(Z \cap X)^{N-3}$ with respect to $(U_2, \dots, U_{N-2})$ a minimizer $\boldsymbol{U}_{\varepsilon\tau}$ satisfies $U_{N-1}^{\varepsilon\tau}, U_N^{\varepsilon\tau} \in Z \cap X$. To see this, note that the final conditions* (6.11c) *yield*

$$\begin{aligned} \alpha_{\varepsilon\tau}^1 U_{N-1}^{\varepsilon\tau} &= \alpha_{\varepsilon\tau}^2 U_{N-2}^{\varepsilon\tau} + \alpha_{\varepsilon\tau}^3 U_{N-3}^{\varepsilon\tau}, \\ \beta_{\varepsilon\tau}^0 U_N^{\varepsilon\tau} &= \beta_{\varepsilon\tau}^1 U_{N-1}^{\varepsilon\tau} + \beta_{\varepsilon\tau}^2 U_{N-2}^{\varepsilon\tau}, \end{aligned}$$

where $\alpha_{\varepsilon\tau}^i, \beta_{\varepsilon\tau}^i$ are suitable constants. Since the right-hand side in the first equation is in $Z \cap X$ so is $U_{N-1}^{\varepsilon\tau}$ and analogously for $U_N^{\varepsilon\tau}$.

Lemma 6.5.3 (Well-posedness of the discrete problem) *For sufficiently small $\varepsilon > 0$ and $\tau > 0$ and all u^0, $u^1 \in H$, $\boldsymbol{W}_{\varepsilon\tau}$ admits a unique minimizer $\boldsymbol{U}_{\varepsilon\tau}$ in $\mathbb{K}_\tau(u^0, u^1)$.*

Proof: This argument is the discrete analogon of the proof of Theorem 6.3.1. In particular, we decompose $\boldsymbol{W}_{\varepsilon\tau}$ into a quadratic part $\boldsymbol{Q}_{\varepsilon\tau}$ and a convex remainder $\boldsymbol{R}_{\varepsilon\tau}$ as

$$\begin{aligned}\boldsymbol{W}_{\varepsilon\tau}[\boldsymbol{U}] &= \frac{\varepsilon^2\rho}{2}\sum_{j=2}^{N}\tau e_\varepsilon^j\|\boldsymbol{\delta}^2U_j\|^2 + \frac{\varepsilon\nu}{2}\sum_{j=2}^{N-1}\tau e_\varepsilon^{j+1}\|\boldsymbol{\delta}U_j\|^2\\ &\quad+\frac{1}{2}\sum_{j=2}^{N-2}\tau e_\varepsilon^{j+2}\Big\{\langle\mathcal{A}U_j,U_j\rangle-\lambda^-\|U_j\|^2\Big\} + \sum_{j=2}^{N-2}\tau e_\varepsilon^{j+2}\mathcal{G}(U_j)\\ &= \boldsymbol{Q}_{\varepsilon\tau}[\boldsymbol{U}] + \boldsymbol{R}_{\varepsilon\tau}[\boldsymbol{U}]\end{aligned}$$

with $\mathcal{G}(U) = \int_\Omega[F(U)-\lambda|U|^2/2]\,\mathrm{d}x$ being convex. The result follows by checking that, for small ε and τ, the functional $\boldsymbol{Q}_{\varepsilon\tau}$ is uniformly convex. To this end, for $\boldsymbol{U}\in\mathbb{K}_\tau(u^0,u^1)$ let $\boldsymbol{V}=(V_0,\dots,V_N)$ be defined as $V_j=(e_\varepsilon^j)^{1/2}U_j$, i.e., $\boldsymbol{V}$ plays the discrete counterpart of the auxiliary function $t\mapsto v(t)=\mathrm{e}^{-t/(2\varepsilon)}u(t)$ in the proof of Theorem 6.3.1. Then, using a discrete product rule we compute

$$\begin{aligned}\sqrt{e_\varepsilon^j}\boldsymbol{\delta}U_j &= \ell_\tau\boldsymbol{\delta}V_j + \frac{1}{2\varepsilon_\tau}V_j, \quad \text{and}\\ \sqrt{e_\varepsilon^j}\boldsymbol{\delta}^2U_j &= \ell_\tau\boldsymbol{\delta}^2V_j + \frac{\ell_\tau}{2\varepsilon_\tau}\boldsymbol{\delta}V_{j-1} + \frac{1}{2\varepsilon_\tau}\boldsymbol{\delta}V_j + \frac{1}{4\varepsilon_\tau^2}V_{j-1},\end{aligned}$$

where $\ell_\tau=\sqrt{\varepsilon/(\varepsilon+\tau)}$ and $\varepsilon_\tau=\tau/(2-2\ell_\tau)$, in particular, we have $\ell_\tau\to1$ and $\varepsilon_\tau\to\varepsilon$ for $\tau\to0$ (compare with the continuous case in (6.13)). Hence, substituting $\boldsymbol{V}$ for $\boldsymbol{U}$ we can rewrite the quadratic part $\boldsymbol{Q}_{\varepsilon\tau}[\boldsymbol{U}]$ as

$$\begin{aligned}\boldsymbol{Q}_{\varepsilon\tau}[\boldsymbol{U}] &= \sum_{j=2}^{N}\frac{\varepsilon^2\rho\tau}{2}\Big\{\ell_\tau^2\|\boldsymbol{\delta}^2V_j\|^2 + \ell_\tau^2\frac{1}{4\varepsilon_\tau^2}\|\boldsymbol{\delta}V_{j-1}\|^2 + \frac{1}{4\varepsilon_\tau^2}\|\boldsymbol{\delta}V_j\|^2 + \frac{1}{16\varepsilon_\tau^4}\|V_{j-1}\|^2\Big\}\\ &\quad+\sum_{j=2}^{N-1}\frac{\varepsilon\nu\tau e_\varepsilon^1}{2}\Big\{\ell_\tau^2\|\boldsymbol{\delta}V_j\|^2 + \frac{1}{4\varepsilon_\tau^2}\|V_{j-1}\|^2\Big\} + \sum_{j=2}^{N-2}\frac{\tau e_\varepsilon^2}{2}\Big\{\langle\mathcal{A}V_j,V_j\rangle - \lambda^-\|V_j\|^2\Big\}\\ &\quad+\boldsymbol{M}_{\varepsilon\tau}[\boldsymbol{V}],\end{aligned}\tag{6.23}$$

where $\boldsymbol{M}_{\varepsilon\tau}[\boldsymbol{V}]$ collects the mixed terms such that we have

$$\begin{aligned}\boldsymbol{M}_{\varepsilon\tau}[\boldsymbol{V}] &= \varepsilon^2\rho\sum_{j=2}^{N}\tau\Big\{\frac{\ell_\tau}{2\varepsilon_\tau}\langle\boldsymbol{\delta}^2V_j,\boldsymbol{\delta}V_j+\ell_\tau\boldsymbol{\delta}V_{j-1}\rangle + \frac{1}{8\varepsilon_\tau^3}\langle\boldsymbol{\delta}V_j+\ell_\tau\boldsymbol{\delta}V_{j-1},V_{j-1}\rangle\Big\}\\ &\quad+\varepsilon^2\rho\sum_{j=2}^{N}\frac{\tau\ell_\tau}{4\varepsilon_\tau^2}\Big\{\langle\boldsymbol{\delta}V_{j-1},\boldsymbol{\delta}V_j\rangle + \langle\boldsymbol{\delta}^2V_j,V_{j-1}\rangle\Big\} + \varepsilon\nu e_\varepsilon^1\sum_{j=2}^{N-1}\frac{\tau\ell_\tau}{2\varepsilon_\tau}\langle\boldsymbol{\delta}V_j,V_j\rangle.\end{aligned}$$

We will treat each of the terms above separately as in the continuous case using summation by parts. Note, however, that due to the discretization additional terms appear which also have to be taken care of. Namely, for the first term in $\boldsymbol{M}_{\varepsilon\tau}[\boldsymbol{V}]$ we have

$$\frac{\varepsilon^2\rho\ell_\tau}{2\varepsilon_\tau}\sum_{j=2}^{N}\tau\langle\boldsymbol{\delta}^2V_j,\boldsymbol{\delta}V_j+\ell_\tau\boldsymbol{\delta}V_{j-1}\rangle = \frac{\varepsilon^2\rho(1-\ell_\tau)\ell_\tau}{4\varepsilon_\tau}\sum_{j=2}^{N}\|\boldsymbol{\delta}V_j-\boldsymbol{\delta}V_{j-1}\|^2 + \frac{\varepsilon^2\rho(1+\ell_\tau)\ell_\tau}{4\varepsilon_\tau}\Big(\|\boldsymbol{\delta}V_N\|^2-\|\boldsymbol{\delta}V_1\|^2\Big).$$

Hence, as $0<\ell_\tau<1$ the first term is convex in $\boldsymbol{V}$. Next, we consider the second term in $\boldsymbol{M}_{\varepsilon\tau}[\boldsymbol{V}]$ and obtain after some rearrangements

$$\frac{\varepsilon^2\rho}{8\varepsilon_\tau^3}\sum_{j=2}^{N}\tau\langle\boldsymbol{\delta}V_j+\ell_\tau\boldsymbol{\delta}V_{j-1},V_{j-1}\rangle = -\frac{\varepsilon^2\rho(1-\ell_\tau)\tau}{16\varepsilon_\tau^3}\sum_{j=2}^{N}\tau\|\boldsymbol{\delta}V_{j-1}\|^2 + \frac{\varepsilon^2\rho}{16\varepsilon_\tau^3}\Big(\ell_\tau\|V_{N-1}\|^2-\ell_\tau\|V_0\|^2+\|V_N\|^2-\|V_1\|^2-\tau^2\|\boldsymbol{\delta}V_N\|^2+\tau^2\|\boldsymbol{\delta}V_1\|^2\Big),$$

where the last two boundary terms are due to shifting the summation index j. The sum in the right-hand side goes with the second term in (6.23). Using the relation $\tau/(2\varepsilon_\tau)=1-\ell_\tau$ we see that the sum of both is positive if $2\ell_\tau>1$, which is true for sufficiently small τ. Next, we compute

$$\frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}\sum_{j=2}^{N}\tau\langle\boldsymbol{\delta}V_j,\boldsymbol{\delta}V_{j-1}\rangle = \frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}\sum_{j=2}^{N}\tau\Big(\|\boldsymbol{\delta}V_j\|^2-\tau\langle\boldsymbol{\delta}^2V_j,\boldsymbol{\delta}V_j\rangle\Big)$$
$$= \frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}\sum_{j=2}^{N}\tau\Big(\|\boldsymbol{\delta}V_j\|^2-\tau^2\|\boldsymbol{\delta}^2V_j\|^2\Big)-\frac{\varepsilon^2\rho\ell_\tau\tau}{4\varepsilon_\tau^2}\Big(\|\boldsymbol{\delta}V_N\|^2-\|\boldsymbol{\delta}V_1\|^2\Big).$$

Using again the relation $\tau/(2\varepsilon_\tau)=1-\ell_\tau$ we can absorb the second term in the sum into the first term in (6.23) provided τ is sufficiently small. For the fourth term in $\boldsymbol{M}_{\varepsilon\tau}$ we apply summation by parts once again which yields the equation

$$\frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}\sum_{j=2}^{N}\tau\langle\boldsymbol{\delta}^2V_j,V_{j-1}\rangle = \frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}(\langle\boldsymbol{\delta}V_N,V_N\rangle-\langle\boldsymbol{\delta}V_1,V_1\rangle)-\frac{\varepsilon^2\rho\ell_\tau}{4\varepsilon_\tau^2}\sum_{j=2}^{N}\tau\|\boldsymbol{\delta}V_j\|^2.$$

Here, the sum in the right-hand side goes with the first term in the last equation.

Finally, for the last sum in $\boldsymbol{M}_{\varepsilon\tau}$ we have

$$\frac{\varepsilon\nu e_\varepsilon^1\ell_\tau}{2\varepsilon_\tau}\sum_{j=2}^{N-1}\tau\langle\boldsymbol{\delta}V_j,V_j\rangle = \frac{\varepsilon\nu e_\varepsilon^1\ell_\tau}{4\varepsilon_\tau}\sum_{j=2}^{N}\|V_j-V_{j-1}\|^2+\frac{\varepsilon\nu e_\varepsilon^1\ell_\tau}{4\varepsilon_\tau}\Big(\|V_N\|^2-\|V_1\|^2\Big).$$

Now, collecting all terms involving V_j we find the quadratic form

$$\boldsymbol{V} \mapsto \sum_{j=3}^{N-1} \frac{\tau\big(\varepsilon^2\rho+4\varepsilon\nu e_\varepsilon^1\varepsilon_\tau^2\ell_\tau-16\varepsilon_\tau^4 e_\varepsilon^2\lambda^-\big)}{32\varepsilon_\tau^4}\|V_{j-1}\|^2$$

which is convex in $\boldsymbol{V}$ if the term in the parentheses is positive, namely if $\ell_\tau\varepsilon_\tau^2/\varepsilon \le \frac{1}{4}\max\{\sqrt{\rho/\lambda^-},\nu/\lambda^-\}$. Obviously, this condition is the discrete analogon of (6.16).

Collecting all boundary terms in a quadratic form, denoted $\boldsymbol{B}_{\varepsilon\tau}$, we argue as in the continuous case and obtain the (uniform) convexity of $\boldsymbol{Q}_{\varepsilon\tau}$. $\square$

6.5.2 Discrete estimate for minimizers of the discrete WIDE functional

The formal argument of Subsection 6.4.1 that led to the crucial estimate for the minimizers u_ε of the continuous WIDE functional $\boldsymbol{W}_\varepsilon$ can be made rigorous at the time-discrete level. Here we present a time-discrete version of estimate (6.18) by using the time-discrete Euler-Lagrange system (6.21a). Namely, we aim at proving the following.

Proposition 6.5.4 (Discrete estimate) *Let $\boldsymbol{U}_{\varepsilon\tau}$ be a stationary point of the discrete* WIDE *functional $\boldsymbol{W}_{\varepsilon\tau}$ in $\mathbb{K}_\tau(u^0,u^1)$. Then, for all ε and τ sufficiently small*

$$\sum_{j=2}^{N-2} \tau\Big\{(\rho+\nu)\|\boldsymbol{\delta} U_j^{\varepsilon\tau}\|^2 + \mathcal{E}(U_j^{\varepsilon\tau})\Big\} \le C, \tag{6.24}$$

where C is constant independent of ε and τ.

Proof: Let us assume from the very beginning that $\rho > 0$ throughout this proof. Indeed, the case $\rho = 0$ (and correspondingly $\nu > 0$) is already treated in [MiS11]. Moreover, let us write $\boldsymbol{U}$ instead of $\boldsymbol{U}_{\varepsilon\tau}$ to keep notation simple. We argue by mimicking the estimate of Subsection 6.4.1 at the discrete level. Namely, we shall perform the following:

$$0 = \sum_{j=2}^{N-2} \tau\,(6.21\text{a})\cdot(\boldsymbol{\delta} U_j - u^1) + \sum_{j=2}^{N-2}\tau\sum_{i=1}^{j}\tau\,(6.21\text{a})\cdot(\boldsymbol{\delta} U_i - u^1). \tag{6.25}$$

At first, let us test the time-discrete Euler–Lagrange equation in (6.21a) by $V_i = \boldsymbol{\delta} U_i - u^1$ and sum for $i = 2,\ldots,j \le N-2$ in order to get that

$$\begin{aligned}\varepsilon^2\rho\sum_{i=2}^{j}\langle\boldsymbol{\delta}^4 U_{i+2},\boldsymbol{\delta} U_i-u^1\rangle - 2\varepsilon\rho\sum_{i=2}^{j}\langle\boldsymbol{\delta}^3 U_{i+1},\boldsymbol{\delta} U_i-u^1\rangle - \varepsilon\nu\sum_{i=2}^{j}\langle\boldsymbol{\delta}^2 U_{i+1},\boldsymbol{\delta} U_i-u^1\rangle \\ + \rho\sum_{i=2}^{j}\langle\boldsymbol{\delta}^2 U_i,\boldsymbol{\delta} U_i-u^1\rangle + \nu\sum_{i=2}^{j}\langle\boldsymbol{\delta} U_i,\boldsymbol{\delta} U_i-u^1\rangle + \sum_{i=2}^{j}\langle \mathrm{D}\mathcal{E}(U_i),\boldsymbol{\delta} U_i-u^1\rangle = 0.\end{aligned} \tag{6.26}$$

We now treat separately all terms in the above left-hand side. In particular, the fourth-order-in-time term can be handled as follows using summation by parts twice and the initial condition $\boldsymbol{\delta} U_1 = u^1$:

$$\varepsilon^2\rho\sum_{i=2}^{j}\tau\langle\boldsymbol{\delta}^4 U_{i+2},\boldsymbol{\delta} U_i-u^1\rangle = \varepsilon^2\rho\langle\boldsymbol{\delta}^3 U_{j+2},\boldsymbol{\delta} U_j-u^1\rangle - \frac{\varepsilon^2\rho}{2}\big\|\boldsymbol{\delta}^2 U_{j+1}\big\|^2 + \frac{\varepsilon^2\rho}{2}\big\|\boldsymbol{\delta}^2 U_2\big\|^2 + \frac{\varepsilon^2\rho}{2}\sum_{i=2}^{j}\big\|\boldsymbol{\delta}^2 U_{i+1}-\boldsymbol{\delta}^2 U_i\big\|^2.$$

Next, we treat the third-order-in-time term of (6.26) somehow similarly using summation by parts once and the initial condition in order to obtain the identity

$$-2\varepsilon\rho\sum_{i=2}^{j}\tau\langle\boldsymbol{\delta}^3 U_{i+1},\boldsymbol{\delta} U_i-u^1\rangle = -2\varepsilon\rho\langle\boldsymbol{\delta}^2 U_{j+1},\boldsymbol{\delta} U_j-u^1\rangle + 2\varepsilon\rho\sum_{i=2}^{j}\tau\big\|\boldsymbol{\delta}^2 U_i\big\|^2.$$

Moreover, we proceed analogously for both of the remaining second-order-in-time derivatives in (6.26) and compute

$$\rho\sum_{i=2}^{j}\tau\langle\boldsymbol{\delta}^2 U_i,\boldsymbol{\delta} U_i-u^1\rangle = \frac{\rho}{2}\big\|\boldsymbol{\delta} U_j-u^1\big\|^2 + \frac{\rho}{2}\sum_{i=2}^{j}\big\|\boldsymbol{\delta} U_i-\boldsymbol{\delta} U_{i-1}\big\|^2, \quad \text{and}$$

$$-\varepsilon\nu\sum_{i=2}^{j}\tau\langle\boldsymbol{\delta}^2 U_{i+1},\boldsymbol{\delta} U_i-u^1\rangle = \frac{\varepsilon\nu}{2}\big\|\boldsymbol{\delta} U_2-u^1\big\|^2 + \frac{\varepsilon\nu}{2}\sum_{i=2}^{j}\big\|\boldsymbol{\delta} U_{i+1}-\boldsymbol{\delta} U_i\big\|^2 - \frac{\varepsilon\nu}{2}\big\|\boldsymbol{\delta} U_{j+1}-u^1\big\|^2.$$

Since the nonlinearity F is assumed to be λ-convex we can use the estimate in (6.9) in order to obtain for the derivative of $\mathcal{E}$ in (6.26)

$$\sum_{i=2}^{j}\tau\langle \mathrm{D}\mathcal{E}(U_i),\boldsymbol{\delta} U_i-u^1\rangle \geq \mathcal{E}(U_j)-\mathcal{E}(u^0) - \sum_{i=1}^{j}\tau\langle \mathrm{D}\mathcal{E}(U_i),u^1\rangle + \frac{\lambda}{2}\sum_{i=1}^{j}\big\|U_i-U_{i-1}\big\|^2,$$

where we also used the initial condition $\boldsymbol{\delta} U_1 = u^1$. Moreover, in the case $\lambda^- \neq 0$ we write $\frac{\lambda^-}{2}\|U_i-U_{i-1}\|^2 = \frac{\lambda^-\tau^2}{2}\big\|\boldsymbol{\delta} U_i\big\|^2$. For sufficiently small τ this term will be absorbed in the remaining terms.

We now recollect the computations above into equation (6.26) in order to deduce the following estimate which holds for $j = 2,\dots,N-2$

$$\begin{aligned}
&\varepsilon^2\rho\langle\boldsymbol{\delta}^3 U_{j+2},\boldsymbol{\delta} U_j-u^1\rangle - \frac{\varepsilon^2\rho}{2}\big\|\boldsymbol{\delta}^2 U_{j+1}\big\|^2 - 2\varepsilon\rho\langle\boldsymbol{\delta}^2 U_{j+1},\boldsymbol{\delta} U_j-u^1\rangle + \mathcal{E}(U_j)\\
&\quad + \frac{\rho}{2}\big\|\boldsymbol{\delta} U_j-u^1\big\|^2 - \frac{\varepsilon\nu}{2}\big\|\boldsymbol{\delta} U_{j+1}-u^1\big\|^2 + \sum_{i=2}^{j}\tau\left\{\nu\big\|\boldsymbol{\delta} U_i\big\|^2 + 2\varepsilon\rho\big\|\boldsymbol{\delta}^2 U_i\big\|^2 - \frac{\lambda^-\tau}{2}\big\|\boldsymbol{\delta} U_i\big\|^2\right\}\\
&\quad \leq C + \sum_{i=1}^{j}\tau\langle \mathrm{D}\mathcal{E}(U_i)+\nu\boldsymbol{\delta} U_i,u^1\rangle,
\end{aligned}\tag{6.27}$$

where C is a constant independent of ε and τ. By choosing $j = N-2$ and taking advantage of the final boundary conditions $\varepsilon^2\rho\boldsymbol{\delta}^3 U_N = \varepsilon\rho\boldsymbol{\delta}^2 U_{N-1}+\varepsilon\nu\boldsymbol{\delta}U_{N-1}$ we arrive at the estimate

$$\begin{aligned}
&\langle \varepsilon\nu\boldsymbol{\delta}U_{N-1} - \varepsilon\rho\boldsymbol{\delta}^2U_{N-1}, \boldsymbol{\delta}U_{N-2}-u^1\rangle - \frac{\varepsilon^2\rho}{2}\|\boldsymbol{\delta}^2U_{N-1}\|^2 + \mathcal{E}(U_{N-2})\\
&+\frac{\rho}{2}\|\boldsymbol{\delta}U_{N-2}-u^1\|^2 - \frac{\varepsilon\nu}{2}\|\boldsymbol{\delta}U_{N-1}-u^1\|^2 + \sum_{i=2}^{N-2}\tau\left\{2\varepsilon\rho\|\boldsymbol{\delta}^2U_i\|^2 + \nu\|\boldsymbol{\delta}U_i\|^2\right\}\\
&\le C + \sum_{i=1}^{N-2}\tau\left\{\langle \mathrm{D}\mathcal{E}(U_i)+\nu\boldsymbol{\delta}U_i, u^1\rangle + \frac{\lambda^-\tau}{2}\|\boldsymbol{\delta}U_i\|^2\right\}.
\end{aligned}\tag{6.28}$$

To treat the boundary terms we use the definition of the difference quotients and rewrite the final conditions in (6.21c) in order to obtain the identity $-\rho\boldsymbol{\delta}^2U_{N-1}=\frac{\tau\nu}{\tau+\varepsilon}\boldsymbol{\delta}U_{N-1}$. Hence, we reformulate the first two terms in (6.28) in terms of $\boldsymbol{\delta}U_{N-1}$ and after some computations obtain the estimate

$$\varepsilon\langle \nu\boldsymbol{\delta}U_{N-1} - \rho\boldsymbol{\delta}^2U_{N-1}, \boldsymbol{\delta}U_{N-2}-u^1\rangle \ge \varepsilon\nu(1-\alpha)\|\boldsymbol{\delta}U_{N-1}\|^2 - \frac{\varepsilon\nu}{\alpha}\|u^1\|^2,$$

where $0 < \alpha < 1$ is an arbitrary constant. Therefore, for sufficiently small $\varepsilon > 0$ we can absorb also the remaining boundary terms such that from (6.28) we arrive at

$$\begin{aligned}
&\frac{\varepsilon\nu}{8}\|\boldsymbol{\delta}U_{N-1}\|^2 + \frac{\rho}{2}\|\boldsymbol{\delta}U_{N-2}-u^1\|^2 + \mathcal{E}(U_{N-2}) + \sum_{i=2}^{N-2}\tau\left\{2\varepsilon\rho\|\boldsymbol{\delta}^2U_i\|^2 + \nu\|\boldsymbol{\delta}U_i\|^2\right\}\\
&\le C + \sum_{i=1}^{N-2}\tau\left\{\langle \mathrm{D}\mathcal{E}(U_i) + \nu\boldsymbol{\delta}U_i, u^1\rangle + \frac{\lambda^-\tau}{2}\|\boldsymbol{\delta}U_i\|^2\right\}.
\end{aligned}\tag{6.29}$$

Let us now move to the consideration of the second term in (6.25). In particular, we multiply the estimate in (6.27) by τ and take the sum for $j = 2,\dots,N-2$ in order to obtain

$$\begin{aligned}
&\varepsilon^2\rho\sum_{j=2}^{N-2}\tau\left\{\langle\boldsymbol{\delta}^3U_{j+2}, \boldsymbol{\delta}U_j-u^1\rangle - \frac{1}{2}\|\boldsymbol{\delta}^2U_{j+1}\|^2\right\} - 2\varepsilon\rho\sum_{j=2}^{N-2}\tau\langle\boldsymbol{\delta}^2U_{j+1}, \boldsymbol{\delta}U_j-u^1\rangle\\
&+\sum_{j=2}^{N-2}\sum_{i=2}^{j}\tau^2\left\{2\varepsilon\rho\|\boldsymbol{\delta}^2U_i\|^2+\nu\|\boldsymbol{\delta}U_i\|^2\right\} + \sum_{j=2}^{N-2}\tau\left\{\frac{\rho}{2}\|\boldsymbol{\delta}U_j-u^1\|^2-\frac{\varepsilon\nu}{2}\|\boldsymbol{\delta}U_{j+1}-u^1\|^2+\mathcal{E}(U_j)\right\}\\
&\le C + \sum_{j=2}^{N-2}\sum_{i=1}^{j}\tau^2\left\{\langle\mathrm{D}\mathcal{E}(U_i)+\nu\boldsymbol{\delta}U_i, u^1\rangle + \frac{\lambda^-\tau}{2}\|\boldsymbol{\delta}U_i\|^2\right\}.
\end{aligned}\tag{6.30}$$

Using again summing by parts and Cauchy's inequality we estimate the first sum in (6.30) in the following way

$$
\begin{aligned}
\varepsilon^2\rho\sum_{j=2}^{N-2}\tau\Big\{\langle\boldsymbol{\delta}^3U_{j+2},\boldsymbol{\delta}U_j-u^1\rangle-\frac{1}{2}\big\|\boldsymbol{\delta}^2U_{j+1}\big\|^2\Big\} &= -\varepsilon^2\rho\sum_{j=2}^{N-2}\tau\Big\{\langle\boldsymbol{\delta}^2U_{j+1},\boldsymbol{\delta}^2U_j\rangle+\frac{1}{2}\big\|\boldsymbol{\delta}^2U_{j+1}\big\|^2\Big\}\\
\geq -\varepsilon^2\rho\sum_{j=2}^{N-2}\Big\{\big\|\boldsymbol{\delta}^2U_{j+1}\big\|^2+\frac{1}{2}\big\|\boldsymbol{\delta}^2U_j\big\|^2\Big\} &\geq -\frac{3\varepsilon^2\rho}{2}\sum_{j=2}^{N-2}\tau\big\|\boldsymbol{\delta}^2U_j\big\|^2-\varepsilon^2\rho\big\|\boldsymbol{\delta}^2U_{N-1}\big\|^2 \qquad (6.31)
\end{aligned}
$$

where we have shifted the indices and used the initial condition $\rho\boldsymbol{\delta}U_1 = \rho u^1$ and the final condition $\rho\boldsymbol{\delta}^2U_N = 0$. The first term in the right-hand side goes together with the corresponding term in (6.29) which is of order ε. Moreover, for the third term in (6.30) we sum by parts in order to obtain the estimate

$$
\begin{aligned}
-2\varepsilon\rho\sum_{j=2}^{N-2}\tau\langle\boldsymbol{\delta}^2U_{j+1},\boldsymbol{\delta}U_j-u^1\rangle &\geq \varepsilon\rho\big\|\boldsymbol{\delta}U_2-u^1\big\|^2-\varepsilon\rho\big\|\boldsymbol{\delta}U_{N-2}-u^1\big\|^2\\
&\quad -2\varepsilon\rho\tau\langle\boldsymbol{\delta}^2U_{N-1},\boldsymbol{\delta}U_{N-2}-u^1\rangle\\
&\geq -\frac{3\varepsilon\rho}{2}\big\|\boldsymbol{\delta}U_{N-2}-u^1\big\|^2-2\varepsilon\rho\tau^2\big\|\boldsymbol{\delta}^2U_{N-1}\big\|^2, \qquad (6.32)
\end{aligned}
$$

where we have used Cauchy's inequality. Note that we have treated the last term in the sum separately in order to be able to absorb the boundary term $\boldsymbol{\delta}U_{N-2}$ into the corresponding term in (6.29). Moreover, the second boundary term $\boldsymbol{\delta}^2U_{N-1}$ can be treated using again the identity $-\rho\boldsymbol{\delta}^2U_{N-1}{=}\frac{\tau\nu}{\tau+\varepsilon}\boldsymbol{\delta}U_{N-1}$.

By taking the sum of (6.29) and (6.30), using the estimates (6.31)—(6.32) and letting ε and τ small enough we obtain that

$$
\begin{aligned}
&\frac{\rho}{4}\big\|\boldsymbol{\delta}U_{N-2}-u^1\big\|^2+\frac{\varepsilon\nu}{4}\big\|\boldsymbol{\delta}U_{N-1}\big\|^2+\mathcal{E}(U_{N-2})+\sum_{j=2}^{N-2}\sum_{i=2}^{j}\tau^2\Big\{2\varepsilon\rho\big\|\boldsymbol{\delta}^2U_i\big\|^2+\nu\big\|\boldsymbol{\delta}U_i\big\|^2\Big\}\\
&+\sum_{j=2}^{N-2}\tau\Big\{\frac{1}{2}\nu\big\|\boldsymbol{\delta}U_j\big\|^2+\frac{\rho}{2}\big\|\boldsymbol{\delta}U_j-u^1\big\|^2+\frac{\rho\varepsilon}{2}\big\|\boldsymbol{\delta}^2U_j\big\|^2+\mathcal{E}(U_j)\Big\}\\
&\leq C+(1+T)\sum_{i=1}^{N-2}\tau\Big\{\big|\langle \mathrm{D}\mathcal{E}(U_i)+\nu\boldsymbol{\delta}U_i,u^1\rangle\big|+\frac{\lambda^-\tau}{2}\big\|\boldsymbol{\delta}U_i\big\|^2\Big\}.
\end{aligned}
$$

As ε and τ are assumed to be small, by using the growth conditions in (6.4) and Young's inequality we readily get the estimate. □

6.5.3 Γ-convergence of discrete WIDE functionals

In order to conclude the proof of Lemma 6.4.1 we need to show that the time-discrete energy estimate in Proposition 6.5.4 passes to the limit as $\tau \to 0$ (for fixed $\varepsilon > 0$). To this aim, we check the discrete-to-continuous Γ-convergence $\boldsymbol{W}_\varepsilon = \Gamma\text{-}\lim_{\tau\to 0} \boldsymbol{W}_{\varepsilon\tau}$ with respect to the weak topology on $\mathbb{V}$ (see [Bra02, Dal93] for relevant definitions and results on Γ-convergence).

For all vectors $\boldsymbol{V} \in H^{N+1}$, we indicate by $\overline{v}_\tau$ and v_τ their backward constant and piecewise affine interpolants on the partition $\{i\tau \,:\, i = 0, \dots, N\}$, respectively. Namely, we have $\overline{v}_\tau(0) = v_\tau(0) = V_0$ and

$$\left.\begin{array}{l} \overline{v}_\tau(t) \equiv V_i, \\ v_\tau(t) = \alpha_i(t) V_i + (1{-}\alpha_i(t)) V_{i-1} \end{array}\right\} \quad \text{for } t \in \,](i{-}1)\tau, i\tau],\ i = 1, \dots, N,$$

where we have used the auxiliary functions

$$\alpha_i(t) = (t - (i{-}1)\tau)/\tau \quad \text{for } t \in \,](i{-}1)\tau, i\tau]\,,\ i = 1, \dots, N.$$

With these definitions we reformulate the estimate in Proposition 6.5.4 as

$$(\rho{+}\nu) \int_\tau^{T-2\tau} \Big(\|u'_{\varepsilon\tau}\|^2 + \mathcal{E}(\overline{u}_{\varepsilon\tau}) \Big)\, \mathrm{d}t \le C, \tag{6.33}$$

where $u_{\varepsilon\tau}$ and $\overline{u}_{\varepsilon\tau}$ denote the piecewise affine and constant interpolants associated with the minimizer $\boldsymbol{U}_{\varepsilon\tau} \in \mathbb{K}_\tau(u^0, u^1)$ of the discrete WIDE functional $\boldsymbol{W}_{\varepsilon\tau}$, respectively.

As a first step in the proof of the Γ-convergence we introduce the space of piecewise affine functions with respect to the partition $\{i\tau \,:\, i = 0, \dots, N\}$ on $[0, T]$ being a subspace of $\mathbb{V}$ and the corresponding convex set $\widehat{\mathbb{K}}_\tau(u^0, u^1)$

$$\begin{aligned} \widehat{\mathbb{V}}_\tau &= \{u : [0, T] \to Z \cap X \,:\, u \text{ is piecewise affine}\} \subset \mathbb{V}, \\ \widehat{\mathbb{K}}_\tau(u^0, u^1) &= \Big\{ u \in \widehat{\mathbb{V}}_\tau \,:\, u(0) = u^0 \text{ and } \rho u \equiv \rho u^1 \text{ on } [0, \tau] \Big\}. \end{aligned}$$

Hence, by identifying the discrete trajectories $\boldsymbol{U} \in \mathbb{Y}_\tau$ with their piecewise affine interpolants $u_\tau \in \widehat{\mathbb{V}}_\tau$ we formulate the minimization of $\boldsymbol{W}_{\varepsilon\tau}$ and $\boldsymbol{W}_\varepsilon$ on the common space $\mathbb{V}$ by extending the WIDE functionals, i.e., we consider

$$\overline{\boldsymbol{W}}_\varepsilon[u] = \begin{cases} \boldsymbol{W}_\varepsilon[u] & \text{if } u \in \mathbb{K}(u^0, u^1), \\ \infty & \text{otherwise}, \end{cases} \qquad \overline{\boldsymbol{W}}_{\varepsilon\tau}[u] = \begin{cases} \boldsymbol{W}_{\varepsilon\tau}[\boldsymbol{U}] & \text{if } u \in \widehat{\mathbb{K}}_\tau(u^0, u^1), \\ \infty & \text{otherwise}, \end{cases}$$

where $\boldsymbol{U} = (u(0), u(\tau), \dots, u(T)) \in \mathbb{Y}_\tau$ for a piecewise affine $u \in \widehat{\mathbb{V}}_\tau$.

As subtle detail note that for an arbitrary $\boldsymbol{U} \in \mathbb{Y}_\tau$ we have in general $U_{N-1}, U_N \notin Z \cap X$ such that the corresponding piecewise affine interpolant u_τ is in general not in $\mathbb{V}$. However, from Remark 6.5.2 we know that the minimizers $\boldsymbol{U}_{\varepsilon\tau}$ of $\mathbb{W}_{\varepsilon\tau}$ satisfy $U^{\varepsilon\tau}_{N-1}, U^{\varepsilon\tau}_N \in Z \cap X$ so that we can neglect this subtlety.

Before we give the main result of this section we note the convergence of the (shifted) interpolants of the time-discrete weights e^i_ε to their continuous counterpart. The proof is being omitted here.

Lemma 6.5.5 *Let $\overline{e}^\varepsilon_\tau$ and e^ε_τ denote the piecewise constant and affine interpolants of the*

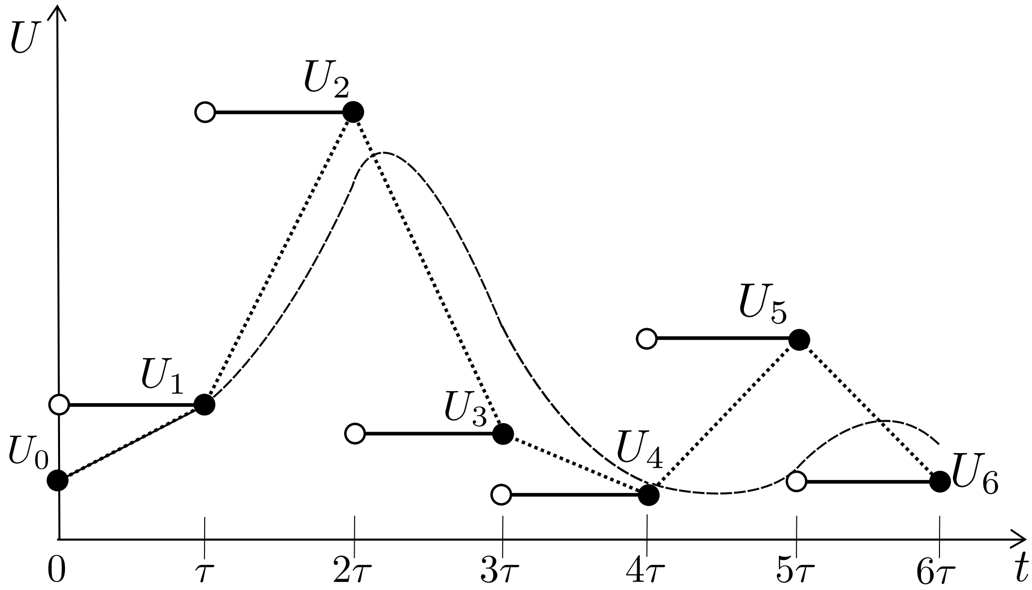

Figure 6.1: Interpolants: piecewise constant (solid), piecewise affine (dotted), piecewise quadratic (dashed)

discrete weights e_ε^i*, respectively. Then*

$$\mathrm{e}_\tau^\varepsilon,\ \overline{e}_\tau^\varepsilon,\ \overline{e}_\tau^\varepsilon(\cdot+\tau),\ \overline{e}_\tau^\varepsilon(\cdot+2\tau) \to \left(t \mapsto \mathrm{e}^{-t/\varepsilon}\right) \textit{ strongly in } \mathrm{L}^\infty(0,T), \tag{6.34}$$

the convergence of $\mathrm{e}_\tau^\varepsilon$ *being actually strong in* $\mathrm{W}^{1,\infty}(0,T)$.

Proposition 6.5.6 (Discrete/continuous Γ-convergence) *The time-discrete WIDE functionals* $\overline{\boldsymbol{W}}_{\varepsilon\tau}$ *converge in the sense of Mosco convergence to the continuous functionals* $\overline{\boldsymbol{W}}_\varepsilon$ *in* $\mathbb{V}$.

Before we prove the Proposition 6.5.6 let us finish the proof of Lemma 6.4.1.

Proof of Lemma 6.4.1: Proposition 6.5.4 yields that the minimizers $u_{\varepsilon\tau}$ of the discrete functional $\overline{\boldsymbol{W}}_{\varepsilon\tau}$ fulfill estimate (6.33) and are hence weakly precompact in $\mathbb{V}$. As $\overline{\boldsymbol{W}}_{\varepsilon\tau}$ Γ-converges to $\overline{\boldsymbol{W}}_\varepsilon$ with respect to the same topology by Proposition 6.5.6 we can apply the Fundamental Theorem of Γ-convergence (see [Dal93, Ch. 7] and [Bra02, Sect. 1.5]), which yields that $u_{\varepsilon\tau} \rightharpoonup u_\varepsilon$ weakly in $\mathbb{V}$, where u_ε is the unique minimizer of $\overline{\boldsymbol{W}}_\varepsilon$. Finally, estimate (6.33) passes to the limit and we have proven Lemma 6.4.1. □

Proof of Proposition 6.5.6: The proof is classically divided into (i) proving the Γ-liminf inequality and (ii) checking the existence of a recovery sequence (see [Dal93, Bra02]).

Ad (i). Assume to be given a sequence $u_\tau \in \widehat{\mathbb{Y}}_\tau$ such that $u_\tau \rightharpoonup u$ with respect to the weak topology on $\mathbb{V}$ and $\liminf_{\tau\to 0} \overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau] < \infty$. Let us denote by $\widetilde{u}_\tau \in \mathrm{H}^2(0,T;Z\cap X)$

the piecewise quadratic interpolant of $U_i = u_\tau(i\tau)$, $i = 0, \ldots, N$, defined by the relations

$$\begin{aligned} \widetilde{u}_\tau(t) &= u_\tau(t) \quad \text{for} \quad t \in [0,\tau] \quad \text{and} \\ \widetilde{u}_\tau'(t) &= \alpha_\tau(t) u_\tau'(t) + (1{-}\alpha_\tau(t)) u_\tau'(t{-}\tau) \quad \text{for} \;\; t \in [\tau, T]\,, \end{aligned}$$

where we have used the notation $\alpha_\tau(t) = \alpha_i(t)$ for $t \in \,](i{-}1)\tau, i\tau]$, $i = 1, \ldots, N$. Hence, $\widetilde{u}_\tau$ is defined such that its derivative is piecewise affine (see Figure 6.1). We preliminarily observe that

$$\widetilde{u}_\tau'(t) = u_\tau'(t{-}\tau) + \tau \alpha_\tau(t) \widetilde{u}_\tau''(t) \quad \text{for almost every } t \in \,]\tau, T]\,. \tag{6.35}$$

Moreover, we check that

$$\overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau] = \int_\tau^T \overline{e}_\tau^\varepsilon \frac{\varepsilon^2 \rho}{2} \|\widetilde{u}_\tau''\|^2 \,\mathrm{d}t + \int_\tau^{T-\tau} \overline{e}_\tau^\varepsilon(\cdot + \tau) \frac{\varepsilon \nu}{2} \|u_\tau'\|^2 \,\mathrm{d}t + \int_\tau^{T-2\tau} \overline{e}_\tau^\varepsilon(\cdot + 2\tau) \mathcal{E}(\overline{u}_\tau) \,\mathrm{d}t.$$

Since by assumption $\liminf_{\tau\to 0} \overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau] < \infty$ we can extract a not relabeled subsequence such that $\limsup_{\tau\to 0} \overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau] < \infty$ and use the convergences of the weights $\overline{e}_\tau^\varepsilon$ in Lemma 6.5.5 to obtain

$$\rho \int_\tau^T \|\widetilde{u}_\tau''\|^2 \,\mathrm{d}t + \nu \int_\tau^{T-\tau} \|u_\tau'\|^2 \,\mathrm{d}t + \int_\tau^{T-2\tau} \mathcal{E}(\overline{u}_\tau) \,\mathrm{d}t \le C.$$

Hence, by using the growth conditions (6.4) and by possibly further extracting a not relabeled subsequence (and considering standard projections for $t > T - 2\tau$) we have the weak convergence of the piecewise constant interpolant

$$\overline{u}_\tau \rightharpoonup u \text{ weakly in } \mathrm{L}^p(0,T;X), \qquad \overline{u}_\tau \rightharpoonup u \text{ weakly in } \mathrm{L}^2(0,T;Z), \tag{6.36}$$

while for the piecewise affine interpolant we have

$$u_\tau \rightharpoonup u \text{ weakly in } \mathrm{H}^1(0,T;H). \tag{6.37}$$

Thus, applying the theorem by Arzelà-Ascoli we even have that $u_\tau \to u$ in $\mathrm{C}(0,T;H)$. In particular, an easy calculation shows that $u_\tau - \overline{u}_\tau \to 0$ in $\mathrm{L}^2(0,T;H)$ such that we arrive at

$$\overline{u}_\tau \to u \quad \text{in } \mathrm{L}^2(0,T;H). \tag{6.38}$$

Furthermore, there exists a v such that for the piecewise quadratic interpolant we obtain

$$\widetilde{u}_\tau \rightharpoonup v \text{ weakly in } \mathrm{H}^2(0,T;H), \quad \rho \widetilde{u}_\tau' \rightharpoonup \rho v' \text{ strongly in } \mathrm{C}(0,T;H). \tag{6.39}$$

Indeed, we have that $v = u$. In order to check this fix $w \in \mathrm{L}^2(0,T;H)$ and compute that

$$\begin{aligned}\int_0^T \langle \widetilde{u}_\tau' - u', w\rangle \,\mathrm{d}t &= \int_0^\tau \langle u_\tau' - u', w\rangle \,\mathrm{d}t + \int_\tau^T \langle u_\tau'(\cdot-\tau) + \tau\alpha_\tau \widetilde{u}_\tau'' - u', w\rangle \,\mathrm{d}t \\ &= \int_0^T \langle u_\tau' - u', w\rangle \,\mathrm{d}t + \int_\tau^T \langle u_\tau'(\cdot-\tau) - u_\tau' + \tau\alpha_\tau \widetilde{u}_\tau'', w\rangle \,\mathrm{d}t \\ &= \int_0^T \langle u_\tau' - u', w\rangle \,\mathrm{d}t - \tau \int_\tau^T (1-\alpha_\tau)\langle \widetilde{u}_\tau'', w\rangle \,\mathrm{d}t \;\longrightarrow\; 0,\end{aligned}$$

where we have used the identity in (6.35), the convergence of the piecewise affine interpolant (6.37), and the boundedness $|\alpha_\tau| \le 1$ and of $\widetilde{u}''$ in $\mathrm{L}^2(0,T;H)$. Hence, we have the convergence $\rho\widetilde{u}_\tau' \rightharpoonup \rho u'$ in $\mathrm{L}^2(0,T;H)$ and $v = u$. In particular, owing to the convergence in (6.39) we have proved that $\rho u^1 = \rho\widetilde{u}_\tau'(0) = \rho u'(0)$ and $u \in \mathbb{K}(u^0,u^1)$.

Eventually, we exploit the strong convergences in $\mathrm{L}^\infty(0,T)$ of the piecewise constant interpolants of the discrete weights in Lemma (6.5.5) and the convergences in (6.36)–(6.39) in order to get by the weak lower semi-continuity of the L^2-norm

$$\begin{aligned}\int_0^T \mathrm{e}^{-t/\varepsilon}\frac{\varepsilon^2\rho}{2}\|u''\|^2 \,\mathrm{d}t &\le \liminf_{\tau\to 0} \int_\tau^T \overline{e}_\tau \frac{\varepsilon^2\rho}{2}\|\widetilde{u}_\tau''\|^2 \,\mathrm{d}t, \\ \int_0^T \mathrm{e}^{-t/\varepsilon}\frac{\varepsilon\nu}{2}\|u'\|^2 \,\mathrm{d}t &\le \liminf_{\tau\to 0} \int_\tau^{T-\tau} \overline{e}_\tau(\cdot+\tau)\frac{\varepsilon\nu}{2}\|u_\tau'\|^2 \,\mathrm{d}t.\end{aligned}$$

Due to (6.38) we can extract a (not relabeled) subsequence such that $\overline{u}_\tau$ converges a.e. in $\Omega \times [0,T]$. Thus, together with $\overline{u}_\tau \rightharpoonup u$ in $\mathrm{L}^2(0,T;Z)$, the application of Fatou's lemma yields

$$\int_0^T \mathrm{e}^{-t/\varepsilon}\mathcal{E}(u)\,\mathrm{d}t \le \liminf_{\tau\to 0} \int_\tau^{T-2\tau} \overline{e}_\tau(\cdot+2\tau)\mathcal{E}(\overline{u}_\tau)\,\mathrm{d}t.$$

In particular, these lower estimates ensure

$$\overline{\boldsymbol{W}}_\varepsilon[u] \le \liminf_{\tau\to 0} \boldsymbol{W}_{\varepsilon\tau}[\boldsymbol{U}] = \liminf_{\tau\to 0} \overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau],$$

which is the desired Γ-$\liminf$ inequality.

Ad (ii). In order to construct a recovery sequence for a given $u \in \mathbb{K}(u^0,u^1)$ we define first the *backward floating mean operator* M_τ on $\mathrm{L}^1(0,T;H)$ (also called *Steklov averaging operator*, see [LSU68, Ch. 2 Sect. 4]) by setting

$$M_\tau[u](t) = \begin{cases} u^0 & \text{for } t \in [0,\tau[, \\ \dfrac{1}{\tau}\displaystyle\int_{t-\tau}^t u(s)\,\mathrm{d}s & \text{for } t \in [\tau,T], \end{cases} \qquad \text{for} \quad u \in \mathrm{L}^1(0,T,H).$$

In particular, using Lebesgue's differentiation theorem we immediately check that for $u \in \mathrm{L}^q(0,T;H)$ (resp. $\mathrm{L}^q(0,T;Z)$, $\mathrm{L}^q(0,T;X)$) we have the convergence $M_\tau[u] \to u$ in $\mathrm{L}^q(0,T;H)$ for $1 \le q < \infty$ (resp. $\mathrm{L}^q(0,T;Z)$, $\mathrm{L}^q(0,T;X)$).

Letting an arbitrary $u \in \mathbb{K}(u^0,u^1)$ be fixed we define the discrete trajectory $\boldsymbol{U} =$

$(U_0, \dots, U_N) \in \mathbb{Y}_\tau$ by

$$U_0 = u^0, \quad \rho U_1 = \rho(u^0 + \tau u^1), \quad U_i = M_\tau[u](i\tau) \quad \text{for } i = 2, \dots, N.$$

We denote by u_τ and $\overline{u}_\tau$ the piecewise affine and constant interpolants, respectively, associated with $\boldsymbol{U}$.

We aim to show that u_τ is a recovery sequence for u. Indeed, we clearly have that $\overline{u}_\tau$ converges strongly to u in $\mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X)$, while u_τ converges at least weakly to u in $\mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X)$. Indeed, we immediately check that for $B = Z, X$ or H and $1 \le q < \infty$ we can estimate

$$\|\overline{u}_\tau\|_{\mathrm{L}^q(\tau,T;B)} \le \|u\|_{\mathrm{L}^q(0,T;B)}, \quad \text{for } u \in \mathrm{L}^q(0,T;B).$$

Moreover, we check that

$$\begin{aligned} \int_0^T \|u_\tau' - u'\|^2 \,\mathrm{d}t &= \int_0^\tau \|u^1 - u'\|^2 \,\mathrm{d}t + \sum_{i=2}^N \int_{(i-1)\tau}^{i\tau} \left\| \frac{1}{\tau^2} \int_{(i-1)\tau}^{i\tau} (u(s) - u(s-\tau)) \,\mathrm{d}s - u' \right\|^2 \mathrm{d}t \\ &= \int_0^\tau \|u^1 - u'\|^2 \,\mathrm{d}t + \sum_{i=2}^N \int_{(i-1)\tau}^{i\tau} \left\| \fint_{(i-1)\tau}^{i\tau} M_\tau[u'](s) \,\mathrm{d}s - u' \right\|^2 \mathrm{d}t. \end{aligned} \tag{6.40}$$

Hence, as one has that $M_\tau[u'] \to u'$ in $\mathrm{L}^2(0,T;H)$ we conclude that $u_\tau \to u$ strongly in $\mathrm{H}^1(0,T;H)$. In particular, we have verified that $u_\tau \rightharpoonup u$ weakly in $\mathbb{V}$.

Next, we exploit the λ-convexity of F and compute that

$$\begin{aligned} &\int_\tau^{T-2\tau} \overline{e}_\tau^\varepsilon(\cdot{+}2\tau) \mathcal{E}(\overline{u}_\tau) \,\mathrm{d}t = \sum_{i=2}^{N-2} \int_{(i-1)\tau}^{i\tau} \left\{ e_\varepsilon^{i+2} (\mathcal{E}(\overline{u}_\tau) - \mathcal{E}(u)) + e_\varepsilon^{i+2} \mathcal{E}(u) \right\} \mathrm{d}t \\ &\le \int_\tau^{T-2\tau} \overline{e}_\tau(\cdot + 2\tau) \left\{ \langle \mathcal{A}\overline{u}_\tau, \overline{u}_\tau {-} u \rangle + \langle \mathcal{B}(\overline{u}_\tau), \overline{u}_\tau {-} u \rangle - \frac{\lambda}{2} \|\overline{u}_\tau {-} u\|^2 + \mathcal{E}(u) \right\} \mathrm{d}t. \end{aligned}$$

In particular, by taking the $\limsup$ as $\tau \to 0$ and recalling that $\overline{u}_\tau \to u$ strongly in $\mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X)$ and the convergences (6.34), we have that

$$\limsup_{\tau \to 0} \int_\tau^{T-2\tau} \overline{e}_\tau^\varepsilon(\cdot{+}2\tau) \mathcal{E}(\overline{u}_\tau) \,\mathrm{d}t \le \int_0^T \mathrm{e}^{-t/\varepsilon} \mathcal{E}(u) \,\mathrm{d}t. \tag{6.41}$$

Next, we deal with the second-order derivatives in time like we did in the first-order case in (6.40). We compute

$$
\begin{aligned}
\rho\int_{3\tau}^{T}\|\widetilde{u}_\tau''-u''\|^2\,\mathrm{d}t &= \rho\sum_{i=4}^{N}\int_{(i-1)\tau}^{i\tau}\left\|\frac{U_i-2U_{i-1}+U_{i-2}}{\tau^2}-u''\right\|^2\mathrm{d}t\\
&= \rho\sum_{i=4}^{N}\int_{(i-1)\tau}^{i\tau}\left\|\frac{1}{\tau^3}\int_{(i-1)\tau}^{i\tau}(u-u(\cdot-\tau))\,\mathrm{d}s-\frac{1}{\tau^3}\int_{(i-2)\tau}^{(i-1)\tau}(u-u(\cdot-\tau))\,\mathrm{d}s-u''\right\|^2\mathrm{d}t\\
&= \rho\sum_{i=4}^{N}\int_{(i-1)\tau}^{i\tau}\left\|\fint_{(i-1)\tau}^{i\tau}M_\tau[M_\tau[u'']]\,\mathrm{d}s-u''\right\|^2\mathrm{d}t\to 0, \qquad (6.42)
\end{aligned}
$$

where the convergence to 0 is ensured by the fact that $M_\tau[M_\tau[u'']]\to u''$ strongly in $\mathrm{L}^2(0,T;H)$. Thus, we have that $\widetilde{u}_\tau''\to u''$ in $\mathrm{L}^2(0,T;H)$.

Finally, combining (6.40)–(6.42) we have proved that

$$
\begin{aligned}
\overline{\boldsymbol{W}}_\varepsilon[u] &= \int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{\varepsilon^2\rho}{2}\|u''\|^2+\frac{\varepsilon\nu}{2}\|u'\|^2+\mathcal{E}(u)\right)\mathrm{d}t\\
&\geq \limsup_{\tau\to 0}\left(\int_\tau^T\overline{e}_\tau\frac{\varepsilon^2\rho}{2}\|\widetilde{u}_\tau''\|^2\,\mathrm{d}t+\int_\tau^{T-\tau}\overline{e}_\tau(\cdot+\tau)\frac{\varepsilon\nu}{2}\|u_\tau'\|^2\,\mathrm{d}t+\int_\tau^{T-2\tau}\overline{e}_\tau(\cdot+2\tau)\mathcal{E}(\overline{u}_\tau)\,\mathrm{d}t\right)\\
&= \limsup_{\tau\to 0}\boldsymbol{W}_{\varepsilon\tau}[\boldsymbol{U}] = \limsup_{\tau\to 0}\overline{\boldsymbol{W}}_{\varepsilon\tau}[u_\tau].
\end{aligned}
$$

Namely, u_τ is a recovery sequence for u. □

Before closing this section let us stress that the obtained results can be adapted in order to encompass more general situations. In particular, we can consider unbounded domains (see [Ste11]) as well as different boundary conditions or the presence of additional source terms with no particular intricacy. Moreover, the WIDE approach can be applied to other classes of dissipative equations. For instance, one could recast the WIDE principle for the *strongly damped* wave equation

$$\rho u''-\nu\Delta u'-\Delta u+f(u)=0,$$

suitably combined with boundary and initial conditions by replacing the dissipative term $\varepsilon\nu\|u'\|^2/2$ with the H^1-seminorm $\varepsilon\nu\|\nabla u'\|^2/2$ in the definition of the functional $\boldsymbol{W}_\varepsilon$.

6.6 Γ-convergence of the WIDE functionals

As already mentioned, a remarkable trait of the WIDE approach is its independence of the character of the equation (5.1) as long as $\rho+\nu>0$. In particular, the WIDE formalism is well-suited in order to describe limiting behaviors in the parameters. First of all, by inspecting the proof of Theorem 6.2.2 it is apparent that stationarity of the WIDE functional pass to limits $\rho\to 0$ and $\nu\to 0$ as well as to joint limits $(\rho,\varepsilon)\to(0,0)$ and $(\nu,\varepsilon)\to(0,0)$. On the other hand, by keeping ε fixed we can argue from a variational viewpoint by addressing the limits $\rho\to 0$ and $\nu\to 0$ within the frame of Γ-convergence.

Let us momentarily modify the notation for the WIDE functional $\boldsymbol{W}_\varepsilon$, the function space $\mathbb{Y}$, and the set $\mathbb{K}(u^0,u^1)$ by highlighting the dependence on the parameters ρ and ν as $\boldsymbol{W}_\varepsilon^{\rho,\nu}$, $\mathbb{Y}^\rho$, and $\mathbb{K}^\rho(u^0,u^1)$, respectively. Moreover, for the sake of notational simplicity we incorporate the constraint $u\in\mathbb{K}^\rho(u^0,u^1)$ directly in the functional by letting

$$\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}[u]=\begin{cases}\boldsymbol{W}_\varepsilon^{\rho,\nu}[u] & \text{for } u\in\mathbb{K}^\rho(u^0,u^1),\\ \infty & \text{otherwise.}\end{cases}$$

We have the following result for the nondissipative and viscous limits $\nu\to0$ and $\rho\to0$, respectively.

Proposition 6.6.1 (Γ-convergence) *The functionals $\boldsymbol{W}_\varepsilon^{\rho,\nu}$ converge in the sense of Mosco for $\nu\to0$ on $\mathbb{Y}$ and for $\rho\to0$ on $\mathbb{V}$, respectively. Namely, it holds that*

$$(i)\quad \overline{\boldsymbol{W}}_\varepsilon^{\rho,0}=\Gamma\text{-}\lim_{\nu\to0}\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}\quad \textit{on both } \mathbb{Y}^\rho \textit{ and } \mathbb{Y}^\rho_{\mathrm{weak}},$$
$$(ii)\quad \overline{\boldsymbol{W}}_\varepsilon^{0,\nu}=\Gamma\text{-}\lim_{\rho\to0}\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}\quad \textit{on both } \mathbb{V} \textit{ and } \mathbb{V}_{\mathrm{weak}}.$$

Proof: **(ad)** (i): The existence of a recovery sequence $u_\nu\to u$ in $\mathbb{Y}$ is immediate by the pointwise convergence $\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}[u]\to\overline{\boldsymbol{W}}_\varepsilon^{\rho,0}[u]$ for $\nu\to0$. The Γ-lim inf inequality follows from the fact that $\overline{\boldsymbol{W}}_{\varepsilon\varepsilon}^{\rho,\nu}\geq\overline{\boldsymbol{W}}_\varepsilon^{\rho,0}$ pointwise and $\overline{\boldsymbol{W}}_\varepsilon^{\rho,0}$ is lower semicontinuous with respect to the weak topology of $\mathbb{Y}^\rho$.

(ad) (ii): The Γ-lim inf inequality is immediate as $\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}\geq\overline{\boldsymbol{W}}_\varepsilon^{0,\nu}$ pointwise and the latter is lower semicontinuous with respect to the weak topology of $\mathbb{V}$. As for the recovery sequence, we shall resort here to some singular perturbation technique (in time). In particular, for any given $u\in\mathbb{K}^0(u^0,u^1)$ and almost every $x\in\Omega$ we can find $t\mapsto v_\rho(x,t)\in\mathrm{H}^1_0(0,T)$ solving weakly

$$v_\rho(x,\cdot)-\sqrt{\rho}v_\rho''(x,\cdot)=u'(x,\cdot)-u^1(x).$$

Then, it is a standard matter to prove that $t\mapsto u_\rho(\cdot,t)=u^0+tu^1+\int_0^t v_\rho(\cdot,s)\,\mathrm{d}s\in\mathbb{K}^\rho(u^0,u^1)$ is such that $u_\rho\to u$ strongly in $\mathbb{V}$ and $\sqrt{\rho}u_\rho''\to0$ strongly in $\mathrm{L}^2(0,T;H)$. We hence have that $\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}[u_\rho]\to\overline{\boldsymbol{W}}_\varepsilon^{0,\nu}[u]$ for $\rho\to0$. □

Let us now check that the latter Γ-convergence result is sufficient in order to prove that, as $\rho\to0$ or $\nu\to0$, (subsequences of) minimizers converge to a minimizer of the corresponding limit functional. To this aim, we just need to check for the precompactness of the minimizers of $\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}$ with respect to the weak $\mathbb{Y}$ or $\mathbb{V}$ topology. Let $u_{\rho,\nu}$ be the minimizer of $\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}$ and define $t\mapsto\widehat{u}(t)=u^0+tu^1\in\mathbb{K}^\rho(u^0,u^1)$ then

$$\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}[u_{\rho,\nu}]\leq\overline{\boldsymbol{W}}_\varepsilon^{\rho,\nu}[\widehat{u}].$$

Hence, using the growth conditions in (6.4) the required precompactness follows.

Before closing this subsection let us stress that the above Γ-limits are taken for ε fixed and record that combined Γ-convergence analyses simultaneously for both parameters and $\varepsilon \to 0$ are presently not available. Additional material on Γ-convergence for WIDE functionals in the parabolic case is however to be found in [AkS11, MiO08, MiS11].

6.7 Improved results for the finite-dimensional case

In this last section we consider the finite-dimensional case, namely we consider trajectories $t \mapsto u(t) \in \mathbb{R}^I$ such that $H = X = Z = \mathbb{R}^I$, for $I \in \mathbb{N}$. In connection to classical mechanics (see [Arn89]) let us denote the state variable by $\boldsymbol{q}$ and the potential by $U \in \mathrm{C}^{1,1}(\mathbb{R}^I)$. The definition of the WIDE functionals is

$$\boldsymbol{W}_\varepsilon[\boldsymbol{q}] = \int_0^T \mathrm{e}^{-t/\varepsilon} \left[\frac{\varepsilon^2 \rho}{2} |\boldsymbol{q}''|^2 + \frac{\varepsilon \nu}{2} |\boldsymbol{q}'|^2 + U(\boldsymbol{q}) \right] \mathrm{d}t. \tag{6.43}$$

The existence and uniqueness of minimizers follows as in the infinite-dimensional case in Theorem 6.3.1. Indeed, note that $U \in \mathrm{C}^{1,1}(\mathbb{R}^I)$ implies the existence of a $\lambda > 0$ such that $\boldsymbol{q} \mapsto U(\boldsymbol{q}) + \lambda/2|\boldsymbol{q}|^2$ is convex.

Moreover, a careful look at the previous sections entails that the results hold in this case under the weaker assumption that $0 \le U \in \mathrm{C}^{1,1}_{\mathrm{loc}}(\mathbb{R}^I)$ and

$$\forall \delta > 0 \quad \exists c_\delta \ge 0 \quad \forall \boldsymbol{q} \in \mathbb{R}^I : \quad |\nabla U(\boldsymbol{q})| \le \delta\big(U(\boldsymbol{q}) + |\boldsymbol{q}|^2\big) + c_\delta. \tag{6.44}$$

This follows for instance for U being the sum of a homogeneous and a subcubic potential. In particular, Lemma 6.4.1 holds and we have for the minimizer $\boldsymbol{q}_\varepsilon$ of $\boldsymbol{W}_\varepsilon$

$$(\rho + \nu) \|\boldsymbol{q}_\varepsilon'\|_{\mathrm{L}^2}^2 \le C. \tag{6.45}$$

In particular, as before we obtain the convergence of the minimizers (in fact, points) $\boldsymbol{q}_\varepsilon$ to a solution of the limit equation

$$\rho \boldsymbol{q}'' + \nu \boldsymbol{q}' + \nabla U(\boldsymbol{q}) = 0. \tag{6.46}$$

Moreover, the convergence result of Theorem 6.2.2 can be refined in order to yield a quantitative rate estimate. Here, for the sake of simplicity we consider only the nondissipative case $\nu = 0$.

Theorem 6.7.1 (Error control) *Let $\boldsymbol{q}_\varepsilon$ minimize $\boldsymbol{W}_\varepsilon$, then*

$$\rho \|\boldsymbol{q} - \boldsymbol{q}_\varepsilon\|_{\mathrm{H}^{1+\eta}} \le c(T) \varepsilon^{(1-\eta)/2} \qquad \textit{for all } \eta \in [0, 1[.$$

Proof: The argument relies on establishing an extra estimate. From bound (6.45) and the local Lipschitz continuity of ∇U we have that $\varepsilon^2 \rho \boldsymbol{q}_\varepsilon'''' - 2\varepsilon \rho \boldsymbol{q}_\varepsilon''' + \rho \boldsymbol{q}_\varepsilon''$ is uniformly bounded in $\mathrm{L}^2(0, T; \mathbb{R}^I)$, depending on T. Hence, by integrating its squared norm we have that

$$
\begin{aligned}
&\varepsilon^4 \int_0^T \rho |\boldsymbol{q}_\varepsilon''''(t)|^2 \,\mathrm{d}t + 4\varepsilon^2 \int_0^T \rho |\boldsymbol{q}_\varepsilon'''(t)|^2 \,\mathrm{d}t + \int_0^T \rho |\boldsymbol{q}_\varepsilon''(t)|^2 \,\mathrm{d}t \\
&\le c(T) + 2\varepsilon^3 \int_0^T \rho \boldsymbol{q}_\varepsilon''''(t) \cdot \boldsymbol{q}_\varepsilon'''(t) \,\mathrm{d}t + 2\varepsilon \int_0^T \rho \boldsymbol{q}_\varepsilon'''(t) \cdot \boldsymbol{q}_\varepsilon''(t) \,\mathrm{d}t - \varepsilon^2 \int_0^T \rho \boldsymbol{q}_\varepsilon''''(t) \cdot \boldsymbol{q}_\varepsilon''(t) \,\mathrm{d}t \\
&= c(T) - \varepsilon^3 \rho |\boldsymbol{q}_\varepsilon'''(0)|^2 - \varepsilon \rho |\boldsymbol{q}_\varepsilon''(0)|^2 + \varepsilon^2 \rho \boldsymbol{q}_\varepsilon'''(0) \cdot \boldsymbol{q}_\varepsilon''(0) + 2\varepsilon^2 \int_0^T \rho |\boldsymbol{q}_\varepsilon'''(t)|^2 \,\mathrm{d}t,
\end{aligned}
$$

where we used the final conditions $\boldsymbol{q}_\varepsilon'''(T) = \boldsymbol{q}_\varepsilon''(T) = 0$. This entails that $\varepsilon^2 \rho^{1/2} \boldsymbol{q}_\varepsilon''''$, $\varepsilon \rho^{1/2} \boldsymbol{q}_\varepsilon'''$, and $\rho^{1/2} \boldsymbol{q}_\varepsilon''$ are bounded in $\mathrm{L}^2(0,T;\mathbb{R}^I)$. Moreover, the Gagliardo-Nirenberg inequality ensures that

$$
\begin{aligned}
&\rho^{1/2} \|\boldsymbol{q}_\varepsilon'''\|_{\mathrm{L}^\infty} \le c(T) (\rho^{1/2} \|\boldsymbol{q}_\varepsilon'''\|_{\mathrm{L}^2} + \rho^{1/2} \|\boldsymbol{q}_\varepsilon'''\|_{\mathrm{L}^2}^{1/2} \|\boldsymbol{q}_\varepsilon''''\|_{\mathrm{L}^2}^{1/2}) \le c(T) \left(\frac{1}{\varepsilon} + \frac{1}{\varepsilon^{3/2}} \right), \\
&\rho^{1/2} \|\boldsymbol{q}_\varepsilon''\|_{\mathrm{L}^\infty} \le c(T) \left(1 + \frac{1}{\varepsilon} \right).
\end{aligned} \tag{6.47}
$$

Take now the difference between the Euler-Lagrange equation for $\boldsymbol{W}_\varepsilon$ and the limit equation (6.46), test it on $\boldsymbol{p}_\varepsilon' = \boldsymbol{q}' - \boldsymbol{q}_\varepsilon'$, and integrate on $(0,t)$ getting

$$
\begin{aligned}
\frac{\rho}{2} |\boldsymbol{p}_\varepsilon'(t)|^2 &= -\varepsilon^2 \int_0^t \rho \boldsymbol{q}_\varepsilon''''(s) \cdot \boldsymbol{p}_\varepsilon'(s) \,\mathrm{d}s + 2\varepsilon \int_0^t \rho \boldsymbol{q}_\varepsilon'''(s) \cdot \boldsymbol{p}_\varepsilon'(s) \,\mathrm{d}s \\
&\quad - \int_0^t (\nabla U(\boldsymbol{q}(s)) - \nabla U(\boldsymbol{q}_\varepsilon(s))) \cdot \boldsymbol{p}_\varepsilon'(s) \,\mathrm{d}s \\
&\le -\varepsilon^2 \rho \boldsymbol{q}_\varepsilon'''(t) \cdot \boldsymbol{p}_\varepsilon'(t) + \varepsilon^2 \int_0^t \rho \boldsymbol{q}_\varepsilon'''(s) \cdot \boldsymbol{p}_\varepsilon''(s) \,\mathrm{d}s + 2\varepsilon \rho \boldsymbol{q}_\varepsilon''(t) \cdot \boldsymbol{p}_\varepsilon'(t) \\
&\quad - 2\varepsilon \int_0^t \rho \boldsymbol{q}_\varepsilon''(s) \cdot \boldsymbol{p}_\varepsilon''(s) \,\mathrm{d}s + c \int_0^t \rho |\boldsymbol{p}_\varepsilon(s)| |\boldsymbol{p}_\varepsilon'(s)| \,\mathrm{d}s \\
&\le c(T)\varepsilon + \frac{\rho}{4} |\boldsymbol{p}_\varepsilon'(t)|^2 + c(T) \int_0^t \rho |\boldsymbol{p}_\varepsilon'(s)|^2 \,\mathrm{d}s,
\end{aligned}
$$

where we used (6.47) in the last inequality. Hence, by means of Gronwall's Lemma we get that $\rho \|\boldsymbol{q}' - \boldsymbol{q}_\varepsilon'\|_{\mathrm{L}^\infty}^2 \le c(T)\varepsilon$. By interpolation [BeL76], for all $\eta \in (0,1)$ we have

$$
\rho \|\boldsymbol{q} - \boldsymbol{q}_\varepsilon\|_{(\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,1}} \le c(T) \|\boldsymbol{q}' - \boldsymbol{q}_\varepsilon'\|_{\mathrm{L}^\infty}^{1-\eta} \|\boldsymbol{q} - \boldsymbol{q}_\varepsilon\|_{\mathrm{H}^2}^{\eta} \le c(T) \varepsilon^{(1-\eta)/2}
$$

(which is stronger than the statement). Eventually, we conclude by noting that

$$
(\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,1} \subset (\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,2} \subset (\mathrm{H}^1,\mathrm{H}^2)_{\eta,2} = \mathrm{H}^{1+\eta}
$$

with continuous injections. □

The conclusions of Theorem 6.7.1 hold unchanged for $\nu > 0$ as long as $\rho > 0$ and the proof is indeed an extension of the proposed one. For $\rho = 0$ (and $\nu > 0$) one resorts in the (necessarily weaker) quantitative convergence result $\nu \|\boldsymbol{q} - \boldsymbol{q}_\varepsilon\|_{\mathrm{H}^\eta} \le c(T) \varepsilon^{(1-\eta)/2}$ for

$\eta \in [0, 1)$.

7 The WIDE principle on the half-line

In this chapter we present the results of [LiS13a] and focus on the infinite time horizon version of the WIDE principle, i.e., $T = \infty$ in the definition of the WIDE functional in (5.2). Also in this case the crucial point is the validity of a priori estimates for the minimizers of the WIDE functional $\boldsymbol{W}_\varepsilon$, however, the methods used here are completely different. Since the WIDE functional is defined on the half-line $\mathbb{R}_+$ we can rescale time by considering $t' = t/\varepsilon$. This leads to an equivalent minimization problem with a weight $\exp(-t')$ instead of $\exp(-t/\varepsilon)$. Moreover, we use time reparametrizations of the scaled minimizers to obtain suitable a priori estimates. Here we follow the argument by Serra & Tilli and hence claim no originality here. It is not clear to us whether the approach of reparametrizing time could also be applied to the finite time horizon case of the previous chapter. While the Euler-Lagrange equation for $\boldsymbol{W}_\varepsilon$ corresponds to the final time horizon case (see (6.11a)) we have no additional final conditions. However, note that u being a function in $\mathrm{H}^1(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathrm{L}^2(\Omega))$ implies the integrability conditions

$$t \mapsto \mathrm{e}^{-t/\varepsilon}\|u\|^2,\ t \mapsto \mathrm{e}^{-t/\varepsilon}\|u'\|^2 \in \mathrm{L}^1(\mathbb{R}_+) \tag{7.1}$$

by virtue of some suitable weighted Poincaré inequality (see Lemma 7.4.1 and [SeT12]). The above integrability conditions play a crucial role in the analysis and are specifically addressed in Subsection 7.2 below. We shall start in Section 7.1 by setting the functional analytic framework and formulate the main result of this chapter. Then, in Section 7.2, we comment on the importance of the integrability conditions by providing an illustrative ODE example. Subsequently in Section 7.3 we discuss the possibility of using the WIDE principle as a selection criterion when the uniqueness of solutions of the limit equation (5.1) is not guaranteed. In the main part of this chapter (Section 7.4) we turn to the proof of the main result, i.e., the convergence of minimizers of the WIDE functional to solutions of (5.1) in the infinite time horizon case. Here, in particular, we adapt the methods by Serra & Tilli for our setting. We conclude the chapter by discussing the WIDE principle in the finite-dimensional case, that is, for a finite-dimensional state space. (This was in fact the setting in [LiS13a]). In particular, we show here the connection between finite and infinite time horizon cases and prove the Γ-convergence of the infinite time WIDE functionals to the finite time ones (for fixed ε).

7.1 Preliminaries and main result

With no loss of generality, hereafter we shall assume the potential $\mathcal{E}$ to be nonnegative. Moreover, let us assume the same growth assumptions as in Section 6.2, namely for all

$u \in \mathbb{R}$, $\xi \in \mathbb{R}^d$ and almost every $x \in \Omega$

$$\frac{1}{C}|u|^p \le F(u) + C, \quad |f(u)|^{p'} \le C(1 + |u|^p), \quad \text{and} \quad \xi \cdot \mathbb{A}(x)\xi \ge \gamma_{\mathbb{A}}|\xi|^2, \tag{7.2}$$

where $p \ge 2$.

Hence, we again arrive at the spaces $H = \mathrm{L}^2(\Omega)$, $X = \mathrm{L}^p(\Omega)$ and $Z = \mathrm{H}^1_0(\Omega)$. Denoting by $\mathrm{d}\sigma_\varepsilon(t) = \mathrm{e}^{-t/\varepsilon}\,\mathrm{d}t$ the weighted measure on $\mathbb{R}_+$ we consider analogously to the previous chapter the spaces $\mathbb{V}_\varepsilon$, $\mathbb{Y}_\varepsilon$ given by

$$\begin{aligned}\mathbb{V}_\varepsilon &= \mathrm{H}^1(\mathbb{R}_+, \mathrm{d}\sigma_\varepsilon; H) \cap \mathrm{L}^2(\mathbb{R}_+, \mathrm{d}\sigma_\varepsilon; Z) \cap \mathrm{L}^p(\mathbb{R}_+, \mathrm{d}\sigma_\varepsilon; X),\\ \mathbb{Y}_\varepsilon &= \left\{u \in \mathbb{V}_\varepsilon \;:\; \rho u \in \mathrm{H}^2(\mathbb{R}_+, \mathrm{d}\sigma_\varepsilon; H)\right\}.\end{aligned}$$

Moreover, the convex set $\mathbb{K}_\varepsilon(u^0, u^1)$ is analogously defined as

$$\mathbb{K}_\varepsilon(u^0, u^1) = \left\{u \in \mathbb{Y}_\varepsilon \;:\; u(0) = u^0, \quad \rho u'(0) = \rho u^1\right\}.$$

The WIDE functional on $\mathbb{K}_\varepsilon(u^0, u^1)$ then reads

$$\boldsymbol{W}_\varepsilon[u] = \int_0^\infty \mathrm{e}^{-t/\varepsilon}\left[\frac{\varepsilon^2\rho}{2}\|u''\|^2 + \frac{\varepsilon\nu}{2}\|u'\|^2 + \mathcal{E}(u)\right]\mathrm{d}t. \tag{7.3}$$

Before going on let us comment on the existence of minimizers of $\boldsymbol{W}_\varepsilon$. In the case of F being convex this follows from the Direct Methods.

Theorem 7.1.1 (Existence of minimizers) *For each $\varepsilon > 0$ there exists a minimizer u_ε of $\boldsymbol{W}_\varepsilon$ in $\mathbb{K}_\varepsilon(u^0, u^1)$.*

Proof: Let $(u_k)_{k\in\mathbb{N}}$ be an infimizing sequence in $\mathbb{K}_\varepsilon(u^0, u^1)$. Due to the growth assumptions on $\mathcal{E}$ we can assume that the sequence $(u_k)_{k\in\mathbb{N}}$ is bounded in $\mathbb{Y}_\varepsilon$ (respectively in $\mathbb{V}_\varepsilon$ for $\rho = 0$). Hence, we can extract a not relabeled subsequence converging weakly in $\mathbb{Y}_\varepsilon$ to a limit u (resp. in $\mathbb{V}_\varepsilon$). Moreover, we can assume that u_k is also converging almost everywhere in $\Omega \times \mathbb{R}_+$. Using Fatou's lemma, we conclude that u is a minimizer. $\square$

We are now in position to give the main result of this chapter:

Theorem 7.1.2 *Assume $\rho+\nu > 0$ and let u_ε minimize $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}_\varepsilon(u^0, u^1)$. Then, there exists a (not relabeled) subsequence such that $u_\varepsilon \overset{*}{\rightharpoonup} u$ in $\mathrm{W}^{1,\infty}(\mathbb{R}_+; H)$ for $\rho > 0$ and weakly in $\mathrm{H}^1(\mathbb{R}_+; H)$ for $\rho = 0$, moreover, $u_\varepsilon \rightharpoonup u$ in $\mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}_+; Z) \cap \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}_+; X)$, where u is a weak solution of*

$$\rho u'' + \nu u' + \mathrm{D}\mathcal{E}(u) = 0 \quad \text{in} \quad \mathbb{R}_+. \tag{7.4}$$

7.2 Integrability conditions at infinity

Before going on we shall explicitly remark the crucial role of the two integrability conditions at infinity (7.1) which are fulfilled by all trajectories u in $\mathbb{K}_\varepsilon$. These conditions correspond to the two *missing* boundary conditions needed in order to complement the Euler-Lagrange equation being of fourth-order. In particular, conditions (7.1) are responsible for the *noncausality* of the problem at all levels $\varepsilon > 0$: The solution u at time t depends on *future*, i.e., its values on (t, ∞). Note however that by taking the limit $\varepsilon \to 0$ causality is eventually restored, see (7.4).

In order to illustrate this remark, let us consider the scalar linear ODE situation of $F(q) = q^2/2$ and $\rho = 1$. In this case, the solution of $\varepsilon^2 q_\varepsilon'''' - 2\varepsilon q_\varepsilon''' + q_\varepsilon'' + q_\varepsilon = 0$ can be computed explicitly as $q_\varepsilon(t) = \sum_{k=1}^4 c_k \exp(\lambda_{\varepsilon,k} t)$ with

$$\lambda_{\varepsilon,1} = \frac{1 - a_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,2} = \frac{1 - b_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,3} = \frac{1 + a_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,4} = \frac{1 + b_\varepsilon}{2\varepsilon}.$$

In the latter a_ε, $b_\varepsilon \in \mathbb{C}$ are chosen in such a way that $a_\varepsilon^2 = 1 - 4\varepsilon i$ and $b_\varepsilon^2 = 1 + 4\varepsilon i$, respectively. By exploiting conditions (7.1) we readily check that, necessarily, $c_3 = c_4 = 0$. Hence, solutions of the Euler-Lagrange equation fulfilling (7.1) are of the form $q(t) = c_1 \exp(\lambda_{\varepsilon,1} t) + c_2 \exp(\lambda_{\varepsilon,2} t)$ and we easily check that $\lambda_{\varepsilon,1} \to i$ and $\lambda_{\varepsilon,2} \to -i$ as $\varepsilon \to 0$. This corresponds to the fact that minimizers of $\boldsymbol{W}_\varepsilon$ in $\mathbb{K}_\varepsilon$ converge to a linear combination of sin and cos, i.e., a solution of $q'' + q = 0$.

7.3 The WIDE principle as a selection criterion

When the growth of $f = F'$ is supercritical the uniqueness of a solution of (5.1) is not guaranteed (e.g. see [Str06]). In this case the WIDE principle may serve as a variational selection criterion. Heuristically, this is related to the specific noncausality of the minimization process for all $\varepsilon > 0$. Indeed, differently from the solutions of the limiting differential problem, the minimizers of $\boldsymbol{W}_\varepsilon$ are allowed in some sense to *peek into the future* and to expend some inertia in order to exploit some possible lower-potential state.

We shall illustrate this fact by a finite-dimensional, scalar example. Fix the initial data to be $q^0 = q^1 = 0$ and choose the potential

$$F(q) = \begin{cases} -8(q^+)^{3/2} & \text{for } q \leq 1, \\ 8((2-q)^+)^{3/2} - 16 & \text{for } q > 1. \end{cases}$$

Note that the potential F is C^1 but not λ-convex at $q = 0$. In particular, F is maximal for $q \leq 0$ and minimal for $q \geq 2$.

The corresponding equation for $\rho = 1$ and $\nu = 0$ reads $q'' = 12\sqrt{q^+}$ which, along with the prescribed initial conditions, admits the trivial solution $q \equiv 0$ as well as a continuum of solutions of the form $t \mapsto ((t-h)^+)^4$ for all $h > 0$. The corresponding WI(D)E functional reads

$$\boldsymbol{W}_\varepsilon[q] = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left[\frac{\varepsilon^2}{2} |q''|^2 + F(q) \right] \mathrm{d}t.$$

For all fixed $\varepsilon > 0$, the Euler-Lagrange equation of $\boldsymbol{W}_\varepsilon$ (along with the initial conditions and integrability conditions (7.1) at $t = \infty$) admits multiple solutions as well. At first,

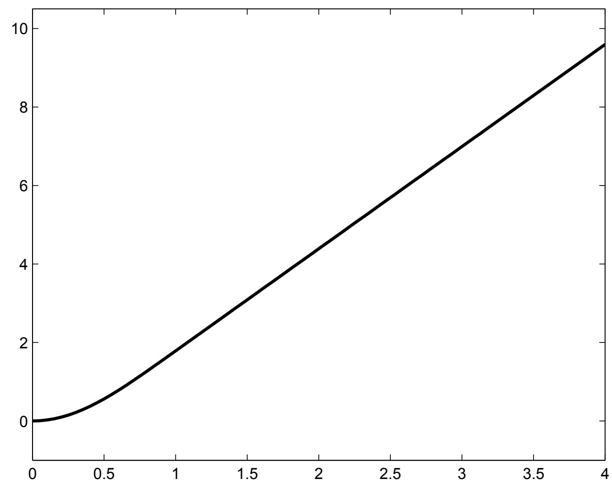

Figure 7.1: The solution q_ε for $\varepsilon = 1/2$.

one has of course the trivial solution. Then, by observing that the potential F is locally Lipschitz continuous for $q > 0$, one can uniquely find the solution q_ε which vanishes just in $t = 0$, see Figure 7.1.

Moreover, as the Euler-Lagrange equation is translation invariant, all trajectories of the form $q_{\varepsilon h}(t) = q_\varepsilon(t-h)$ are solutions as well.

Note that for small times (approximately $t < 1$) we have that $q_\varepsilon'' \neq 0$ and $F(q_\varepsilon)$ is negative but still not minimal. Then, at later times the trajectory q_ε reaches the region where F is minimal and gets basically affine ($q_\varepsilon'' \sim 0$). In particular, the integrand of the WIE functional over q_ε changes sign over time and we can (numerically) evaluate the value $\boldsymbol{W}_\varepsilon[q_\varepsilon]$ to be negative, see Figure 7.2.

As clearly $\boldsymbol{W}_\varepsilon[0] = 0$ for the trivial solution and $\boldsymbol{W}_\varepsilon[q_{\varepsilon h}] = \mathrm{e}^{-h/\varepsilon}\boldsymbol{W}_\varepsilon[q_\varepsilon] > \boldsymbol{W}_\varepsilon[q_\varepsilon]$, one has that the WIE principle selects exactly the trajectory w_ε. Eventually, by taking the limit $\varepsilon \to 0$, the minimizers of the WIE functional can hence be expected to converge to the particular solution $t \mapsto t^4$ of the limiting problem.

7.4 A priori estimate and limit passage

As for the finite time horizon case, the convergence proof of Theorem 7.1.2 follows from an a priori estimate. For this, we will make use of the following Poincaré-type lemma, which can be found in [SeT12, Lemma 2.3].

Lemma 7.4.1 *Let* $w : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ *be a function such that for every* $T > 0$, $w \in$

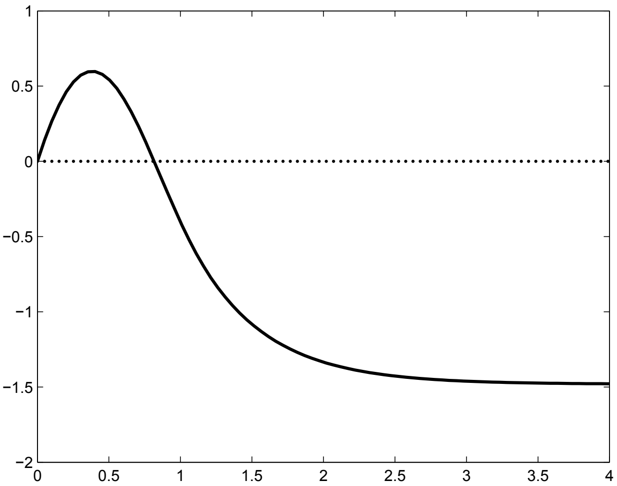

Figure 7.2: The function $t \mapsto \int_0^t \mathrm{e}^{-s/\varepsilon}\left(\frac{\varepsilon^2}{2}|q_\varepsilon''(s)|^2 + F(q_\varepsilon(s))\right)\,\mathrm{d}s$ for $\varepsilon = 1/2$.

$\mathrm{L}^2([0,T];H)$ *and* $w' \in \mathrm{L}^2([0,T];H)$, *then*

$$\int_0^\infty \mathrm{e}^{-s}\|w(s)\|^2\,\mathrm{d}s \leq 2\|w(0)\|^2 + 4\int_0^\infty \mathrm{e}^{-s}\|w'(s)\|^2\,\mathrm{d}s. \tag{7.5}$$

Note that none of the integrals appearing in (7.5) is claimed to be finite.

Lemma 7.4.2 (A priori estimate) *For* $\varepsilon > 0$ *let* u_ε *minimize* $\boldsymbol{W}_\varepsilon$ *on* $\mathbb{K}_\varepsilon(u^0, u^1)$. *Then, there exists a constant* $C > 0$, *independent of* ε, *such that for all* $t \geq 0$

$$\rho\left\|u_\varepsilon'(t)\right\|^2 + \nu\int_0^t \left\|u_\varepsilon'(s)\right\|^2\,\mathrm{d}s \leq C. \tag{7.6}$$

Moreover, for arbitrary $T \geq \varepsilon$, *it holds that*

$$\int_0^T \left(\|u_\varepsilon\|_Z^2 + \|u_\varepsilon\|_X^p\right)\mathrm{d}t \leq CT. \tag{7.7}$$

Before proceeding to the proof, let us remark that the two terms in estimate (7.6) are exactly the ones which are expected in the limit $\varepsilon = 0$. As such, the estimate shows a remarkable optimality with respect to possibly mixed dissipative/nondissipative dynamics. The proof of estimate (7.6) results by extending the argument from [SeT12] in order to handle dissipative effects.

Proof: By letting u_ε be a minimizer of $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}_\varepsilon(u^0, u^1)$ we define for $s = t/\varepsilon$ the rescaled quantities

$$v_\varepsilon(s) \stackrel{\text{def}}{=} u_\varepsilon(\varepsilon s), \quad \boldsymbol{G}_\varepsilon[v] \stackrel{\text{def}}{=} \int_0^\infty \mathrm{e}^{-s}\left(\frac{\rho}{2}\|v''\|^2 + \frac{\varepsilon\nu}{2}\|v'\|^2 + \varepsilon^2\mathcal{E}(v)\right)\mathrm{d}s.$$

Obviously, we have the implication $u_\varepsilon \in \mathbb{Y}_\varepsilon$ iff $v_\varepsilon \in \mathbb{Y}_1$. Furthermore, it holds that $\boldsymbol{W}_\varepsilon[u_\varepsilon] = \boldsymbol{G}_\varepsilon[v_\varepsilon]/\varepsilon$, hence, v_ε is a minimizer of $\boldsymbol{G}_\varepsilon$ on $\mathbb{K}_1(u^0, \varepsilon u^1)$.

By choosing the competitor $\widehat{v}_\varepsilon(s) = u^0 + \arctan(\varepsilon u^1 s)$ (which, in particular, is such that $\widehat{v}_\varepsilon \in \mathbb{K}_1(u^0, \varepsilon u^1)$) due to the minimality of v_ε we have that

$$\boldsymbol{G}_\varepsilon[v_\varepsilon] \le \boldsymbol{G}_\varepsilon[\widehat{v}_\varepsilon] \le \int_0^\infty \mathrm{e}^{-s}\Big(\varepsilon^2 \mathcal{E}(\widehat{v}_\varepsilon) + c(\varepsilon^6\rho + \varepsilon^3\nu)\Big)\,\mathrm{d}s \le c\varepsilon^2.$$

In particular, using Lemma 7.4.1 we can estimate the weighted L^2-norm of v_ε' in terms of the initial condition εu^1 and the weighted L^2-norm of v_ε'', i.e.,

$$\begin{aligned}(\rho+\varepsilon\nu)\int_0^\infty \mathrm{e}^{-s}\|v_\varepsilon'\|^2\,\mathrm{d}s &\le 2\varepsilon^2\|u^1\|^2 + \int_0^\infty \mathrm{e}^{-s}\Big(4\rho\|v_\varepsilon''\|^2 + \varepsilon\nu\|v_\varepsilon'\|^2\Big)\,\mathrm{d}s \\ &\le c\big(\varepsilon^2 + \boldsymbol{G}_\varepsilon[v_\varepsilon]\big) \le c\varepsilon^2,\end{aligned} \tag{7.8}$$

where we used that $\mathcal{E}$ is assumed to be nonnegative. Next, we define the auxiliary function $\boldsymbol{H}_\varepsilon : [0,\infty[\, \to \mathbb{R}$ via

$$\boldsymbol{H}_\varepsilon(r) = \int_r^\infty \mathrm{e}^{-s}\left(\frac{\rho}{2}\big\|v_\varepsilon''\big\|^2 + \frac{\varepsilon\nu}{2}\big\|v_\varepsilon'\big\|^2 + \varepsilon^2\mathcal{E}(v_\varepsilon)\right)\mathrm{d}s.$$

Note that $\boldsymbol{H}_\varepsilon$ is continuous, nonnegative and nonincreasing, and $\boldsymbol{H}_\varepsilon(0) = \boldsymbol{G}_\varepsilon[v_\varepsilon]$. Moreover, Lemma 7.4.3 yields the identity

$$\frac{\rho}{2}\langle v_\varepsilon'', v_\varepsilon'\rangle + \varepsilon\nu\|v_\varepsilon'\|^2 - \frac{\mathrm{d}}{\mathrm{d}r}\Big(\frac{\rho}{2}\langle v_\varepsilon'', v_\varepsilon'\rangle - \frac{1}{2}\mathrm{e}^r\boldsymbol{H}_\varepsilon\Big) = -\rho\|v_\varepsilon''\|^2 \tag{7.9}$$

which holds in the distributional sense. Hence, by defining the function $\boldsymbol{I}_\varepsilon : [0,\infty[\, \to \mathbb{R}$ to be the primitive of the left-hand side of (7.9), i.e.,

$$\boldsymbol{I}_\varepsilon(r) = \frac{\rho}{4}\big\|v_\varepsilon'(r)\big\|^2 + \varepsilon\nu\int_0^r \big\|v_\varepsilon'(s)\big\|^2\,\mathrm{d}s - \frac{\rho}{2}\langle v_\varepsilon''(r), v_\varepsilon'(r)\rangle + \frac{1}{2}\mathrm{e}^r\boldsymbol{H}_\varepsilon(r),$$

we can rewrite the equation in (7.9) in the form

$$\boldsymbol{I}_\varepsilon'(r) = -\rho\big\|v_\varepsilon''(r)\big\|^2. \tag{7.10}$$

In particular, the function $r \mapsto \boldsymbol{I}_\varepsilon(r)$ is nonincreasing and belongs to $\mathrm{W}^{1,1}(0,T)$ for every $T > 0$. Hence, multiplying $\boldsymbol{I}_\varepsilon$ with $r \mapsto \mathrm{e}^{-r}$ and integrating on $]t,T[$ one concludes that

$$\begin{aligned}&\frac{\rho}{4}\mathrm{e}^{-t}\big\|v_\varepsilon'(t)\big\|^2 - \frac{\rho}{4}\mathrm{e}^{-T}\big\|v_\varepsilon'(T)\big\|^2 + \frac{1}{2}\int_t^T \boldsymbol{H}_\varepsilon(r)\,\mathrm{d}r + \varepsilon\nu\int_t^T \mathrm{e}^{-r}\left(\int_0^r \big\|v_\varepsilon'(s)\big\|^2\,\mathrm{d}s\right)\mathrm{d}r \\ &= \int_t^T \mathrm{e}^{-r}\boldsymbol{I}_\varepsilon(r)\,\mathrm{d}r \le (\mathrm{e}^{-t} - \mathrm{e}^{-T})\boldsymbol{I}_\varepsilon(t) \le (\mathrm{e}^{-t} - \mathrm{e}^{-T})\boldsymbol{I}_\varepsilon(0).\end{aligned} \tag{7.11}$$

Let us now take the limit for $T \to \infty$. By recalling that $\mathrm{e}^{-T}\|v_\varepsilon'(T)\|^2 \to 0$ we get

$$\frac{\rho}{4}\mathrm{e}^{-t}\|v_\varepsilon'(t)\|^2 + \varepsilon\nu\int_t^\infty \mathrm{e}^{-s}\left(\int_0^s \|v_\varepsilon'(r)\|^2\,\mathrm{d}r\right)\mathrm{d}s \leq \mathrm{e}^{-t}\boldsymbol{I}_\varepsilon(0).$$

In particular, $t \mapsto \mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s \in \mathrm{L}^1(\mathbb{R}_+)$ and, owing also to bound (7.8), it is a standard matter to compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s\right) = -\mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s + \mathrm{e}^{-t}\|v_\varepsilon'(t)\|^2$$

and deduce that indeed $t \mapsto \mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s \in \mathrm{W}^{1,1}(\mathbb{R}_+)$. Hence, we also have that $\mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s \to 0$ as t goes to ∞.

We shall now go back to relation (7.11), handle the $\varepsilon\nu$-term using integration by parts

$$\varepsilon\nu\int_t^T \mathrm{e}^{-s}\left(\int_0^s \|v_\varepsilon'(r)\|^2\,\mathrm{d}r\right)\mathrm{d}s = -\varepsilon\nu\mathrm{e}^{-T}\int_0^T \|v_\varepsilon'(s)\|^2\,\mathrm{d}s + \varepsilon\nu\mathrm{e}^{-t}\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s \\ + \varepsilon\nu\int_t^T \mathrm{e}^{-s}\|v_\varepsilon'(s)\|^2\,\mathrm{d}s,$$

and take the limit $T \to \infty$ in order to get

$$\frac{\rho}{4}\|v_\varepsilon'(t)\|^2 + \varepsilon\nu\int_0^t \|v_\varepsilon'(s)\|^2\,\mathrm{d}s \leq \boldsymbol{I}_\varepsilon(0).$$

Thus, once we have shown that $\boldsymbol{I}_\varepsilon(0) \leq C\varepsilon^2$ holds, we obtain the estimate in (7.6) by rescaling time. For this note that from (7.8) we infer

$$\int_0^1 \langle v_\varepsilon'', v_\varepsilon'\rangle\,\mathrm{d}s \leq c\varepsilon^2 \quad \text{and} \quad \boldsymbol{H}_\varepsilon(t) \leq \boldsymbol{H}_\varepsilon(0) = \boldsymbol{G}_\varepsilon[v_\varepsilon] \leq c\varepsilon^2.$$

Therefore, we arrive at $\int_0^1 \boldsymbol{I}_\varepsilon(r)\,\mathrm{d}r \leq c\varepsilon^2$ and using (7.10) we compute for almost every $r \in\,]0,1[$

$$\boldsymbol{I}_\varepsilon(0) = \boldsymbol{I}_\varepsilon(r) + \rho\int_0^r \left\|v_\varepsilon''(s)\right\|^2\mathrm{d}s \leq \boldsymbol{I}_\varepsilon(r) + \rho\int_0^1 \left\|v_\varepsilon''(s)\right\|^2\mathrm{d}s.$$

Integrating this estimate over the interval $]0,1[$ finally gives

$$\boldsymbol{I}_\varepsilon(0) \leq \int_0^1 \left(\boldsymbol{I}_\varepsilon(r) + \rho\left\|v_\varepsilon''(r)\right\|^2\right)\mathrm{d}r \leq c\varepsilon^2 \tag{7.12}$$

which proves the estimate in (7.6).

For the second estimate in (7.7) note that due to the growth conditions in (7.2) we have for every $r \geq 0$ the estimate

$$\begin{aligned}\varepsilon^2 \int_r^{r+1} \left(\|v_\varepsilon(s)\|_Z^2 + \|v_\varepsilon(s)\|_X^p\right) \mathrm{d}s &\leq \varepsilon^2 \mathrm{e}^{r+1} \int_r^{r+1} \mathrm{e}^{-s} \left(\|v_\varepsilon(s)\|_Z^2 + \|v_\varepsilon(s)\|_X^p\right) \mathrm{d}s \\ &\leq c\left(\varepsilon^2 + \mathrm{e}^{r+1}\boldsymbol{H}_\varepsilon(r)\right).\end{aligned}$$

Since $r \mapsto \boldsymbol{H}_\varepsilon(r)$ is nonincreasing and positive, we obtain the estimate

$$\frac{\mathrm{e}^r}{2}\boldsymbol{H}_\varepsilon(r{+}1) \leq \frac{\mathrm{e}^r}{2}\int_r^{r+1} \boldsymbol{H}_\varepsilon(s)\,\mathrm{d}s \leq \frac{\mathrm{e}^r}{2}\int_r^\infty \boldsymbol{H}_\varepsilon(s)\,\mathrm{d}s \leq \boldsymbol{I}_\varepsilon(r) \leq c\varepsilon^2,$$

where we used (7.11) and (7.12) for the last two inequalities, respectively. Hence, it holds that $\mathrm{e}^r\boldsymbol{H}_\varepsilon(r) \leq c\varepsilon^2$ for all $r \geq 1$. On the other hand for $r \in [0,1]$ we have $\mathrm{e}^{r-1}\boldsymbol{H}_\varepsilon(r) \leq \boldsymbol{H}_\varepsilon(0) = \boldsymbol{G}_\varepsilon[v_\varepsilon] \leq c\varepsilon^2$. In conclusion we obtain $\mathrm{e}^r\boldsymbol{H}_\varepsilon(r) \leq c\varepsilon^2$ for all $r \geq 0$ and hence

$$\int_r^{r+1} \left(\|v_\varepsilon(s)\|_Z^2 + \|v_\varepsilon(s)\|_X^p\right) \mathrm{d}s < C.$$

Thus, substituting $\eta = \varepsilon s$ we arrive at the following estimate for the minimizer u_ε for all $t > 0$

$$\int_{\varepsilon t}^{\varepsilon t+\varepsilon} \left(\|u_\varepsilon\|_Z^2 + \|u_\varepsilon\|_X^p\right) \mathrm{d}\eta \leq C\varepsilon. \tag{7.13}$$

Fixing now $T \geq \varepsilon$ we can cover the interval $[t, t+T]$ by $k_T^\varepsilon = \lceil T/\varepsilon \rceil$ adjacent subintervals whose length is bounded by ε. On each of these subintervals $[\tau, \tau+\varepsilon]$ we can use (7.13) and then, summing the resulting estimates, we find that

$$\int_t^{t+T} \left(\|u_\varepsilon\|_Z^2 + \|u_\varepsilon\|_X^p\right) \mathrm{d}\eta \leq C\varepsilon k_T^\varepsilon \leq CT,$$

which yields (7.7) as a particular case when $t = 0$ □

In the following lemma we prove the crucial identity for the auxiliary function $\boldsymbol{H}_\varepsilon$ introduced in the proof of Lemma (7.4.2).

Lemma 7.4.3 *Let v_ε minimize the rescaled WIDE functional $\boldsymbol{G}_\varepsilon$ and let $\boldsymbol{H}_\varepsilon : [0,\infty[\to \mathbb{R}$ be as in the proof of Lemma 7.4.2. Then,*

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\rho}{2}\langle v_\varepsilon'', v_\varepsilon'\rangle\right) = \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{1}{2}\mathrm{e}^r\boldsymbol{H}(r)\right) + \rho\|v_\varepsilon''\|^2 + \frac{\rho}{2}\langle v_\varepsilon'', v_\varepsilon'\rangle + \varepsilon\nu\|v_\varepsilon'\|^2.$$

in the distributional sense.

Proof: We fix an arbitrary $\eta \in \mathrm{C}_c^\infty(]0,\infty[)$ and consider for $\delta \in \mathbb{R}$ the function

$$\phi_\delta(s) = s - \delta\int_0^s \eta(r)\,\mathrm{d}r = s - \delta g(s).$$

For $|\delta|$ small enough ϕ_δ is a C^∞-diffeomorphism on $\mathbb{R}_+$. In particular, for sufficiently small $s > 0$ we have $\phi_\delta(s) \equiv s$ such that the auxiliary function $\tilde{v}_\varepsilon^\delta(s) = v_\varepsilon(\phi_\delta(s))$ satisfies $\tilde{v}_\varepsilon^\delta \in \mathbb{K}_1(u^0, \varepsilon u^1)$. Let $\psi_\delta = \phi_\delta^{-1}$ denote the inverse of ϕ_δ, we have by change of variables $s = \psi_\delta(r)$

$$\boldsymbol{G}_\varepsilon[\widetilde{v}_\varepsilon^\delta] = \int_0^\infty \psi_\delta'\,\mathrm{e}^{-\psi_\delta}\left[\frac{\rho}{2}\Big\||\phi_\delta'(\psi_\delta)|^2 v_\varepsilon'' + \phi_\delta''(\psi_\delta)v_\varepsilon'\Big\|^2 + \frac{\varepsilon\nu}{2}\|\phi_\delta'(\psi_\delta)v_\varepsilon'\|^2 + \varepsilon^2\mathcal{E}(v_\varepsilon)\right]\mathrm{d}r.$$

In particular, note that due to the identity $\psi_\delta(r) = r + \delta g(\psi_\delta(r))$ we have that $\psi_\delta(r) \geq r - \delta\|g\|_\infty$ and hence $\mathrm{e}^{-\psi_\delta(r)} \leq \mathrm{e}^{\delta\|g\|_\infty}\mathrm{e}^{-r}$ and $\boldsymbol{G}_\varepsilon[\widetilde{v}_\varepsilon^\delta]$ is finite.

Since $\widetilde{v}_\varepsilon^\delta$ reduces to v_ε when $\delta = 0$, the minimality of v_ε entails

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\boldsymbol{G}_\varepsilon[\widetilde{v}_\varepsilon^\delta]\,|_{\delta=0} = 0. \tag{7.14}$$

To compute this derivative observe that differentiating of ψ_δ with respect to δ leads to the formula

$$\partial_\delta\psi_\delta(r) = g(\psi_\delta) + \delta g'(\psi_\delta(r))\partial_\delta\psi_\delta(r).$$

Since $\psi_0(r) = r$ we obtain $\partial_\delta\psi_\delta(r)|_{\delta=0} = g(r)$. Similarly, differentiation with respect to r yields $\psi_\delta'(r) = 1 + \delta g'(\psi_\delta(r))\psi_\delta'(r)$ such that we obtain

$$\partial_\delta\psi_\delta' = \psi_\delta(r)(g'\circ\psi_\delta) + \delta\big(\psi_\delta'\,\partial_\delta\psi_\delta(g''\circ\psi_\delta) + \partial_\delta\psi_\delta'(g'\circ\psi_\delta)\big).$$

In particular, we obtain $\partial_\delta\psi_\delta'(r)|_{\delta=0} = g'(r)$ and as a consequence we have

$$\partial_\delta\big(\psi_\delta'(r)\mathrm{e}^{-\psi_\delta(r)}\big) = \big(g'(r) - g(r)\big)\mathrm{e}^{-r}.$$

Finally, simple calculations lead to the formulas

$$\frac{\mathrm{d}}{\mathrm{d}\delta}|\phi_\delta'\circ\psi_\delta|^2|_{\delta=0} = -2g', \quad \frac{\mathrm{d}}{\mathrm{d}\delta}\phi_\delta''\circ\psi_\delta|_{\delta=0} = -g'', \quad \frac{\mathrm{d}}{\mathrm{d}\delta}\phi_\delta'\circ\psi_\delta = -g'.$$

Denoting by $L_\varepsilon(\delta, r)$ the function within the square brackets in $\boldsymbol{G}_\varepsilon[\widetilde{v}_\varepsilon^\delta]$ there holds

$$L_\varepsilon(0, r) = \frac{\rho}{2}\|v_\varepsilon''\|^2 + \frac{\varepsilon\nu}{2}\|v_\varepsilon'\|^2 + \varepsilon^2\mathcal{E}(v_\varepsilon).$$

Moreover, using the formulas above we compute

$$\partial_\delta L_\varepsilon(0, r) = -\rho\langle v_\varepsilon'', 2g'(r)v_\varepsilon'' + g''(r)v_\varepsilon'\rangle - \varepsilon\nu\langle v_\varepsilon', g'(r)v_\varepsilon'\rangle.$$

Hence, by combining both identities we see that (7.14) reduces to

$$\int_0^\infty \mathrm{e}^{-r}(g' - g)L_\varepsilon(0)\,\mathrm{d}r = \int_0^\infty \mathrm{e}^{-r}\left(2\rho g'\|v_\varepsilon''\|^2 + \rho g''\langle v_\varepsilon'', v_\varepsilon'\rangle + \varepsilon\nu g'\|v_\varepsilon'\|^2\right)\mathrm{d}r$$

Integrating by parts the term involving $g(r)\mathrm{e}^{-r}L_\varepsilon(0,r) = -g(r)\boldsymbol{H}'_\varepsilon(r)$ gives

$$-\int_0^\infty \eta(\boldsymbol{H}'_\varepsilon + \boldsymbol{H}_\varepsilon)\,\mathrm{d}r = \int_0^\infty \mathrm{e}^{-r}\left(2\rho\eta\|v''_\varepsilon\|^2 + \rho\eta'\langle v''_\varepsilon, v'_\varepsilon\rangle + \varepsilon\nu\eta\|v'_\varepsilon\|^2\right)\mathrm{d}r.$$

where we used that the boundary terms are vanishing due to $g(t) \equiv 0$ for t small and $\boldsymbol{H}_\varepsilon(t) \to 0$ for $t \to \infty$. Since $r \mapsto \mathrm{e}^{-r}\eta(r)$ is an arbitrary test function in $\mathrm{C}_c^\infty(]0,\infty[)$ we obtain the desired identity. $\square$

7.4.1 Proof of the main result

We are now in position to prove the main result of this chapter, namely, the convergence of the minimizers u_ε of the WIDE functional $\boldsymbol{W}_\varepsilon$ to a solution of the limit equation (7.4). The proof is analogous to the finite time horizon case and rests upon the possibility to pass to the limit in the (weak form of the) Euler-Lagrange equation.

From the estimates in Lemma 7.4.2 we see that for every $T > 0$ we can extract a (not relabeled) subsequence such that

$$u_\varepsilon \rightharpoonup u \text{ in } \mathrm{H}^1(0,T;H) \cap \mathrm{L}^2(0,T;Z) \cap \mathrm{L}^p(0,T;X).$$

In particular, we have $u_\varepsilon \to u$ in $\mathrm{C}([0,T];H)$ and hence also $u_\varepsilon \to u$ and $f(u_\varepsilon) \to f(u)$ pointwise almost everywhere. Moreover, we argue as in [Vis96, Prop. 3.10] to obtain $u_\varepsilon \to u$ in $\mathrm{L}^q([0,T]\times\Omega)$ for $1 \le q < p$ and therefore, using (7.2)

$$f(u_\varepsilon) \to f(u) \quad \text{in} \quad \mathrm{L}^r(0, T\times\Omega), \qquad r \in [1,p'[\,. \tag{7.15}$$

As in Section 6.4.2 we take an arbitrary $w \in \mathrm{C}_c^\infty([0,T[;Z\cap X)$ and define the test function $v_\varepsilon \overset{\text{def}}{=} t \mapsto \mathrm{e}^{t/\varepsilon}w(t) - t(w'(0) + \frac{1}{\varepsilon}w(0)) - w(0)$. Hence, testing with v_ε in the weak form of the Euler-Lagrange equation and integrating by parts we use (7.15) and pass to the limit to obtain

$$\int_0^T \left(\langle \mathrm{D}\mathcal{E}(u) + \nu u', w\rangle - \rho\langle u', w'\rangle\right)\mathrm{d}t = -\langle u^1, w(0)\rangle,$$

namely that u is a weak solution of the limit equation in (7.4) with initial conditions $u(0) = u^1$ and $u'(0) = u^1$. If u is the unique solution, the whole sequence converges.

7.5 The finite-dimensional case

As in the discussion of the finite time horizon case in Chapter 6 let us comment here on the finite-dimensional case, which was originally discussed in [LiS13a]. In order to distinguish this case from the infinite-dimensional one we choose the notation $\boldsymbol{q} \in \mathbb{R}^I$, $I \in \mathbb{N}$, for the state of the system and denote by $U \in \mathrm{C}^1(\mathbb{R})$ the potential. Hence, the WIDE functional on $\mathrm{H}^2(\mathbb{R}_+; \mathrm{d}\sigma_\varepsilon; \mathbb{R}^I)$ reads

$$\boldsymbol{W}_\varepsilon[\boldsymbol{q}] = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left[\frac{\varepsilon^2 \rho}{2} |\boldsymbol{q}''|^2 + \frac{\varepsilon \nu}{2} |\boldsymbol{q}'|^2 + U(\boldsymbol{q}) \right] \mathrm{d}t. \tag{7.16}$$

Let us additionally assume that the potential U is bounded from below. Then, by arguing as in the last section we derive the a priori estimate

$$\forall t > 0: \quad \rho |\boldsymbol{q}_\varepsilon'(t)|^2 + \nu \int_0^t |\boldsymbol{q}_\varepsilon'|^2 \,\mathrm{d}s \le c$$

for the minimizers of $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}_\varepsilon(\boldsymbol{q}^0, \boldsymbol{q}^1)$ (defined as before) and obtain the following result.

Theorem 7.5.1 (WIDE principle) *Assume $\rho + \nu > 0$ and let $\boldsymbol{q}_\varepsilon$ minimize the functional $\boldsymbol{W}_\varepsilon$ on $\mathbb{K}_\varepsilon(\boldsymbol{q}^0, \boldsymbol{q}^1)$. Then, for some subsequence $\boldsymbol{q}_{\varepsilon_k}$ we have $\boldsymbol{q}_{\varepsilon_k} \to \boldsymbol{q}$ weakly-$*$ in $\mathrm{W}^{1,\infty}(\mathbb{R}_+; \mathbb{R}^I)$ if $\rho > 0$ and weakly in $\mathrm{H}^1(\mathbb{R}_+; \mathbb{R}^I)$ if $\rho = 0$ (hence, locally uniformly), where*

$$\rho \boldsymbol{q}'' + \nu \boldsymbol{q}' + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \ \text{ in } \ \mathbb{R}_+, \quad \boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \rho \boldsymbol{q}'(0) = \rho \boldsymbol{q}^1.$$

Let us explicitly mention that the latter result holds more generally for two symmetric and positive-definite mass and viscosity matrices M and N such that $M + N > 0$.

Moreover, by inspecting the proof of Lemma 7.4.2 one realizes that the statement of Theorem 7.5.1 is indeed valid in some greater generality. In particular, one could require the potential U to be defined just on a non-empty open subset $D \subset \mathbb{R}^I$ and, by letting U_{ext} be its trivial extension to ∞ out of D, impose

$$0 \le U \in C^1(D) \ \text{ and } \ U_{\mathrm{ext}} \ \text{ be lower semicontinuous.} \tag{7.17}$$

Note that the lower semicontinuity of U_{ext} expresses the fact that the potential U is actually confining the evolution to D. In particular U becomes unbounded by approaching the boundary of D. By requiring $\boldsymbol{q}^0 \in D$, under assumption (7.17) Theorem 7.5.1 still holds. The extension of the WIDE principle to the latter type of potentials is not at all academical as it qualifies the WIDE functional to be applicable also in some singular potential situation.

We shall also mention that, although completely neglected in this text for the sake of simplicity, a suitably well-behaved time-dependence of the potential U (hence, in particular, a non-homogeneous flow) can be considered.

7.5.1 Infinite-horizon Γ-limit

In this final section we comment on the connection between the final time horizon case in Section 6.7 and the infinite time horizon case discussed here. More precisely, denoting the WIDE functionals considered in (6.43) by $\boldsymbol{W}_\varepsilon^T$, where T is the final time, we shall show the Γ-convergence $\boldsymbol{W}_\varepsilon^T \xrightarrow{\Gamma} \boldsymbol{W}_\varepsilon^\infty$, where $\boldsymbol{W}_\varepsilon^\infty$ denotes the functional defined in (7.16).

Let us denote by $\mathbb{Y}_T$ and $\mathbb{Y}_\infty$ the spaces $\mathrm{H}^2(0,T;\mathbb{R}^I)$ and $\mathrm{H}^2(\mathbb{R}_+;\,\mathrm{d}\sigma_\varepsilon;\mathbb{R}^I)$, respectively. We shall embed the space $\mathbb{Y}_T$ into $\mathbb{Y}_\varepsilon$ by identifying a given $\boldsymbol{q}\in\mathbb{Y}_T$ with the unique function $\overline{\boldsymbol{q}}\in\mathbb{Y}_\infty$ satisfying $\overline{\boldsymbol{q}}\equiv\boldsymbol{q}$ on $[0,T]$ and being affine on $]T,\infty[$. Thus, we can extend the functionals $\boldsymbol{W}_\varepsilon^T$ on the common space $\mathbb{Y}_\infty$ by defining

$$\overline{\boldsymbol{W}}_\varepsilon^T[\overline{\boldsymbol{q}}]=\begin{cases}\boldsymbol{W}_\varepsilon^T[\boldsymbol{q}] & \text{if } \overline{\boldsymbol{q}}\in\mathbb{K}_\varepsilon(\boldsymbol{q}^0,\boldsymbol{q}^1) \text{ and } \overline{\boldsymbol{q}} \text{ is affine on } [T,\infty[,\\ +\infty & \text{otherwise.}\end{cases}$$

In particular, if $\overline{\boldsymbol{q}}\in\mathbb{K}_\varepsilon(u^0,u^1)$ minimizes $\overline{\boldsymbol{W}}_\varepsilon^T$ then $\boldsymbol{q}$ is also a minimizer of $\boldsymbol{W}_\varepsilon^T$.

Proposition 7.5.2 (Γ-limit for $T\to\infty$) *Assume that U is quadratically bounded, then $\overline{\boldsymbol{W}}_\varepsilon^T\xrightarrow{\Gamma}\boldsymbol{W}_\varepsilon^\infty$ weakly in $\mathrm{L}^2(\mathbb{R}_+;\,\mathrm{d}\sigma_\varepsilon;\mathbb{R}^I)$. Moreover, minimizers of $\overline{\boldsymbol{W}}_\varepsilon^T$ converge weakly in $\mathrm{L}^2(\mathbb{R}_+;\,\mathrm{d}\sigma_\varepsilon;\mathbb{R}^I)$ (up to subsequences) to minimizers of $\boldsymbol{W}_\varepsilon^\infty$.*

Proof: A recovery sequence for a given $\overline{\boldsymbol{q}}\in\mathbb{K}_\varepsilon(\boldsymbol{q}^0,\boldsymbol{q}^1)$ is easily constructed by defining $\overline{\boldsymbol{q}}_T=\overline{\boldsymbol{q}}$ on $[0,T]$ and $\overline{\boldsymbol{q}}_T$ affine on $[T,\infty[$. Then, it is easy to check that $\overline{\boldsymbol{W}}_\varepsilon^T[\overline{\boldsymbol{q}}_T]\to\boldsymbol{W}_\varepsilon^\infty[\overline{\boldsymbol{q}}]$.

Assume now to be given a sequence $\overline{\boldsymbol{q}}_T$ such that $\overline{\boldsymbol{q}}_T\to\overline{\boldsymbol{q}}$ weakly in $\mathrm{L}^2(\mathbb{R}_+,\,\mathrm{d}\sigma_\varepsilon;\mathbb{R}^I)$. By taking with no loss of generality $\liminf_{T\to\infty}\overline{\boldsymbol{W}}_\varepsilon^T[\overline{\boldsymbol{q}}_T]<\infty$ and using that $\overline{\boldsymbol{q}}_T''=0$ on $]T,\infty[$ we have that

$$\liminf_{T\to\infty}\int_0^T\mathrm{e}^{-t/\varepsilon}|\boldsymbol{q}_T''(t)|^2\,\mathrm{d}t=\liminf_{T\to\infty}\int_0^\infty\mathrm{e}^{-t/\varepsilon}|\boldsymbol{q}_T''(t)|^2\,\mathrm{d}t\geq\int_0^\infty\mathrm{e}^{-t/\varepsilon}|\boldsymbol{q}''(t)|^2\,\mathrm{d}t$$

and $\overline{\boldsymbol{q}}_T\to\overline{\boldsymbol{q}}$ pointwise almost everywhere. Eventually, $\liminf_{T\to0}\overline{\boldsymbol{W}}_\varepsilon^T[\overline{\boldsymbol{q}}_T]\geq\boldsymbol{W}_\varepsilon^\infty[\overline{\boldsymbol{q}}]$ by Dominated Convergence as $U(\overline{\boldsymbol{q}}_T)\leq c(1+|\overline{\boldsymbol{q}}_T|^2)$ (using also Lemma 7.4.1).

Let now $\tilde{\boldsymbol{q}}(t)=\boldsymbol{q}^0+t\boldsymbol{q}^1$. Then all minimizers $\overline{\boldsymbol{q}}_T$ of $\overline{\boldsymbol{W}}_\varepsilon^T$ fulfill

$$\frac{\rho}{2}\int_0^\infty\mathrm{e}^{-t/\varepsilon}|\overline{\boldsymbol{q}}_T''|^2\,\mathrm{d}t=\frac{\rho}{2}\int_0^T\mathrm{e}^{-t/\varepsilon}|\boldsymbol{q}_T''|^2\,\mathrm{d}t\leq\overline{\boldsymbol{W}}_\varepsilon^T[\overline{\boldsymbol{q}}_T]\leq\overline{\boldsymbol{W}}_\varepsilon^T[\tilde{\boldsymbol{q}}]<\infty$$

independently of T. In particular, $\overline{\boldsymbol{q}}_T$ are weakly precompact in $\mathrm{L}^2(\mathbb{R}_+,\,\mathrm{d}\sigma_\varepsilon;\mathbb{R}^I)$. Hence, it converge up to subsequences to a minimizer of $\boldsymbol{W}_\varepsilon^\infty$. $\square$

Bibliography

[AD*11] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Comm. Math. Phys.*, 307(3), 791–815, 2011.

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[AkS10] G. Akagi and U. Stefanelli. A variational principle for doubly nonlinear evolution. *Appl. Math. Lett.*, 23(9), 1120–1124, 2010.

[AkS11] G. Akagi and U. Stefanelli. Weighted energy-dissipation functionals for doubly nonlinear evolution. *J. Funct. Anal.*, 260(9), 2541–2578, 2011.

[AkS12] G. Akagi and U. Stefanelli. Doubly nonlinear evolution equations as convex minimization problems. 2012. In preparation.

[AM*01] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26(1-2), 43–100, 2001.

[AM*12] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: from diffusion to reaction. *Calc. Var. Part. Diff. Eqns.*, 44, 419–454, 2012.

[AmS08] L. Ambrosio and S. Serfaty. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11), 1495–1539, 2008.

[Arn89] V. I. Arnol′d. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[ASZ09] L. Ambrosio, G. Savaré, and L. Zambotti. Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Related Fields*, 145(3-4), 517–564, 2009.

[Att84] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[BaÉ85] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[Bak94] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[Bas07] J.-L. Basdevant. *Variational principles in physics.* Springer, New York, 2007.

[BD*10] M. Burger, M. Di Franscesco, J.-F. Pietschmann, and B. Schlake. Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Analysis*, 42(6), 2842–2871, 2010.

[BeB00] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3), 375–393, 2000.

[BeL76] J. Bergh and J. Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[Ber09] V. L. Berdichevsky. *Variational principles of continuum mechanics. I.* Interaction of Mechanics and Mathematics. Springer-Verlag, Berlin, 2009. Fundamentals.

[BFG06] G. Bellettini, G. Fusco, and N. Guglielmi. A concept of solution and numerical experiments for forward-backward diffusion equations. *Discrete Contin. Dyn. Syst.*, 16(4), 783–842, 2006.

[Bio55] M. A. Biot. Variational principles in irreversible thermodynamics with application to viscoelasticity. *Phys. Rev. (2)*, 97, 1463–1469, 1955.

[BoP11] D. Bothe and M. Pierre. The instantaneous limit for reaction-diffusion systems with a fast irreversible reaction. *Discr. Cont. Dynam. Systems Ser. S*, 8(1), 49–59, 2011.

[Bra02] A. Braides. *Γ-convergence for beginners.* Oxford Lecture Series in Mathematics and its Applications 22. Oxford: Oxford University Press. xii, 2002.

[Bré71] H. Brézis. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In *Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971)*, pages 101–156. Academic Press, New York, 1971.

[Bré73] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[BrE76a] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(20), Ai, A1197–A1198, 1976.

[BrE76b] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17), Aii, A971–A974, 1976.

[CFP06] R. Chill, E. Fašangová, and J. Prüss. Convergence to steady states of solutions of the Cahn-Hilliard and Caginalp equations with dynamic boundary conditions. *Math. Nachr.*, 279(13-14), 1448–1462, 2006.

[CGM08] L. Cherfils, S. Gatti, and A. Miranville. Existence of global solutions to the Caginalp phase-field system with dynamic boundary conditions and singular potentials. *Journal of Mathematical Analysis and Applications*, 343(1), 557 – 566, 2008.

[Cia00] P. G. Ciarlet. *Mathematical elasticity. Vol. III*, volume 29 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam, 2000. Theory of shells.

[CL*10] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4), 1273–1309, 2010.

[Cla90] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[CoO08] S. Conti and M. Ortiz. Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids*, 56(5), 1885–1904, 2008.

[CoR90] P. Colli and J.-F. Rodrigues. Diffusion through thin layers with high specific heat. *Asymptotic Anal.*, 3(3), 249–263, 1990.

[CrP69] M. G. Crandall and A. Pazy. Semi-groups of nonlinear contractions and dissipative sets. *J. Functional Analysis*, 3, 376–418, 1969.

[Dal93] G. Dal Maso. *An Introduction to Γ-Convergence.* Birkhäuser Boston Inc., Boston, MA, 1993.

[DaS08] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Analysis*, 40, 1104–1122, 2008.

[DaS10] S. Daneri and G. Savaré. Lecture notes on gradient flows and optimal transport. arXiv:1009.3737v1, 2010.

[dCa76] M. P. a. do Carmo. *Differential geometry of curves and surfaces.* Englewood Cliffs, N. J.: Prentice-Hall, 1976.

[De 96] E. De Giorgi. Conjectures concerning some evolution problems. *Duke Math. J.*, 81(2), 255–268, 1996.

[DeF06] L. Desvillettes and K. Fellner. Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations. *J. Math. Anal. Appl.*, 319(1), 157–176, 2006.

[DeF07] L. Desvillettes and K. Fellner. Entropy methods for reaction-diffusion systems. In *Discrete Contin. Dyn. Syst. (suppl). Dynamical Systems and Differential*

Equations. Proceedings of the 6th AIMS International Conference, pages 304–312, 2007.

[DeF08] L. Desvillettes and K. Fellner. Entropy methods for reaction-diffusion equations with degenerate diffusion arising in reversibly chemistry. In *Proceedings of EQUADIFF 2007*, 2008. To appear.

[DeM84] S. De Groot and P. Mazur. *Non-Equilibrium Thermodynamics.* Dover Publ., New York, 1984.

[DF*07] L. Desvillettes, K. Fellner, M. Pierre, and J. Vovelle. Global existence for quadratic systems of reaction-diffusion. *Adv. Nonlinear Stud.*, 7(3), 491–511, 2007.

[DMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3), 180–187, 1980.

[ElS10] C. M. Elliott and B. Stinner. Modeling and computation of two phase geometric biomembranes using surface finite elements. *J. Comput. Phys.*, 229(18), 6585–6612, 2010.

[ErM12] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Rational Mech. Anal.*, 206(3), 997–1038, 2012.

[FiG10] A. Figalli and N. Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *J. Math. Pures Appl.*, 94(2), 107–130, 2010.

[FRG*06] A. Favini, G. Ruiz Goldstein, J. A. Goldstein, and S. Romanelli. The heat equation with nonlinear general Wentzell boundary condition. *Adv. Differ. Equ.*, 11(5), 481–510, 2006.

[GaG86] H. Gajewski and K. Gröger. On the basic equations for carrier transport in semiconductors. *J. Math. Anal. Appl.*, 113(1), 12–35, 1986.

[GaG05] H. Gajewski and K. Gärtner. On a nonlocal model of image segmentation. *Z. angew. Math. Mech. (ZAMM)*, 56(4), 572–591, 2005.

[GGM08] C. Gal, M. Grasselli, and A. Miranville. Nonisothermal Allen-Cahn equations with coupled dynamic boundary conditions. Colli, P. (ed.) et al., Proceedings of international conference on: Nonlinear phenomena with energy dissipation. Mathematical analysis, modeling and simulation, Chiba, Japan, November 26–30, 2007. Tokyo: Gakkotosha. Gakuto International Series Mathematical Sciences and Applications 29, 117-139, 2008.

[Gho09] N. Ghoussoub. *Self-dual partial differential systems and their variational principles.* Springer Monographs in Mathematics. Springer, New York, 2009.

[GiM04] V. Giovangigli and M. Massot. Entropic structure of multicomponent reactive flows with partial equilibrium reduced chemistry. *Math. Methods Appl. Sci. (MMAS)*, 27(7), 739–768, 2004.

[GlG09] A. Glitzky and K. Gärtner. Energy estimates for continuous and discretized electro-reaction-diffusion systems. *Nonlinear Anal.*, 70(2), 788–805, 2009.

[GlH05] A. Glitzky and R. Hünlich. Global existence result for pair diffusion models. *SIAM J. Math. Analysis*, 36(4), 1200–1225 (electronic), 2005.

[Gli08] A. Glitzky. Exponential decay of the free energy for discretized electro-reaction-diffusion systems. *Nonlinearity*, 21(9), 1989–2009, 2008.

[Gli09] A. Glitzky. Energy estimates for electro-reaction-diffusion systems with partly fast kinetics. *Discr. Cont. Dynam. Systems Ser. A*, 25(1), 159–174, 2009.

[Gli12] A. Glitzky. An electronic model for solar cells including active interfaces and energy resolved defect densities. *SIAM J. Math. Analysis*, 44, 3874–3900, 2012.

[GlM13] A. Glitzky and A. Mielke. A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. *Z. angew. Math. Phys. (ZAMP)*, 64(1), 29–52, 2013.

[Gri85] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[Gri04] J. A. Griepentrog. On the unique solvability of a nonlocal phase separation problem for multicomponent systems. In *Nonlocal elliptic and parabolic problems*, volume 66 of *Banach Center Publ.*, pages 153–164. Polish Acad. Sci., Warsaw, 2004.

[GST09] U. Gianazza, G. Savaré, and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Rational Mech. Anal.*, 194(1), 133–220, 2009.

[Gur63] M. E. Gurtin. Variational principles in the linear theory of viscoelasticity. *Arch. Rational Mech. Anal.*, 13, 179–191, 1963.

[Gur64a] M. E. Gurtin. Variational principles for linear elastodynamics. *Arch. Rational Mech. Anal.*, 16, 34–50, 1964.

[Gur64b] M. Gurtin. Variational principles for linear initial-value problems. *Q. Appl. Math.*, 22, 252–256, 1964.

[Hel09] E. Hellinger. Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. *J. Reine Angew. Math.*, 136, 210–271, 1909.

[Hla69] I. Hlaváček. Variational principles for parabolic equations. *Apl. Mat.*, 14, 278–297, 1969.

[Ilm94] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108(520), x+90, 1994.

[JKO98] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Analysis*, 29(1), 1–17, 1998.

[Kak48] S. Kakutani. On equivalence of infinite product measures. *Ann. of Math. (2)*, 49, 214–224, 1948.

[KE*01] R. Kenzler, F. Eurich, P. Maass, B. Rinn, J. Schropp, E. Bohl, and W. Dieterich. Phase separation in confined geometries: Solving the Cahn-Hilliard equation with generic boundary conditions. *Computer Physics Communications*, 133(2-3), 139 – 157, 2001.

[KjB08] S. Kjelstrup and D. Bedeaux. *Non-equilibrium thermodynamics of heterogeneous systems*, volume 16 of *Series on Advances in Statistical Mechanics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[KMM06] M. Kurzke, C. Melcher, and R. Moser. Domain walls and vortices in thin ferromagnetic films. In *Analysis, modeling and simulation of multiscale problems*, pages 249–298. Springer, Berlin, 2006.

[Kom67] Y. Komura. Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan*, 19, 493–507, 1967.

[Kra95] G. Krausch. Surface induced self assembly in thin polymer films. *Materials Science and Engineering: R: Reports*, 14(1-2), v – 94, 1995.

[Kur07] M. Kurzke. The gradient flow motion of boundary vortices. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 24(1), 91–112, 2007.

[Lán70] C. Lánczos. *The variational principles of mechanics*. Mathematical Expositions, No. 4. University of Toronto Press, Toronto, Ont., fourth edition, 1970.

[Lax06] P. D. Lax. *Hyperbolic partial differential equations*, volume 14 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2006. With an appendix by Cathleen S. Morawetz.

[Lie12] M. Liero. Passing from bulk to bulk/surface evolution in the Allen-Cahn equation. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 2012. to appear.

[LiM72] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[LiM12] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. Royal Soc. A*, 2012. to appear.

[Lio69] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.

[Lis09] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3), 712–740, 2009.

[LiS13a] M. Liero and U. Stefanelli. A new minimum principle for Lagrangian mechanics. *J. Nonlinear Sci.*, 2013. to appear.

[LiS13b] M. Liero and U. Stefanelli. Weighted inertia-dissipation-energy functionals for semilinear equations. *Boll. Unione Mat. Ital.*, 9, 2013. to appear.

[LiV87] F. Liese and I. Vajda. *Convex statistical distances*, volume 95 of *Teubner-Texte zur Mathematik*. BSB B. G. Teubner Verlagsgesellschaft, 1987.

[LMN08] M. Lucia, C. B. Muratov, and M. Novaga. Existence of traveling waves of invasion for Ginzburg-Landau-type problems in infinite cylinders. *Arch. Ration. Mech. Anal.*, 188(3), 475–508, 2008.

[LMS12] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Diff. Eqns.*, 253(2), 814–850, 2012.

[LOR09] C. J. Larsen, M. Ortiz, and C. L. Richardson. Fracture paths from front kinetics: relaxation and rate independence. *Arch. Ration. Mech. Anal.*, 193(3), 539–583, 2009.

[LSU68] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*. Transl. Math. Monographs. Amer. Math. Soc., Providence, R.I., 1968.

[Maa11] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261, 2250–2292, 2011.

[McC97] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128, 153–179, 1997.

[Mie08] A. Mielke. Weak-convergence methods for Hamiltonian multiscale problems. *Discrete Contin. Dyn. Syst.*, 20(1), 53–79, 2008.

[Mie11a] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *J. Mathem. Pures Appl.*, 2011. Submitted. WIAS preprint 1650.

[Mie11b] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24, 1329–1346, 2011.

[Mie13] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discr. Cont. Dynam. Systems Ser. S*, 6(2), 479–499, 2013.

[MiO08] A. Mielke and M. Ortiz. A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems. *ESAIM Control Optim. Calc. Var.*, 14(3), 494–516, 2008.

[MiS08] A. Mielke and U. Stefanelli. A discrete variational principle for rate-independent evolution. *Adv. Calc. Var.*, 1(4), 399–431, 2008.

[MiS11] A. Mielke and U. Stefanelli. Weighted energy-dissipation functionals for gradient flows. *ESAIM Control Optim. Calc. Var.*, 17(1), 52–85, 2011.

[MiZ05] A. Miranville and S. Zelik. Exponential attractors for the Cahn-Hilliard equation with dynamic boundary conditions. *Mathematical Methods in the Applied Sciences*, 28(6), 709–735, 2005.

[MMS09] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12), 1352–1397, 2009.

[Moi04] B. L. Moiseiwitsch. *Variational principles.* Dover Publications Inc., Mineola, NY, 2004. Corrected reprint of the 1966 original.

[MOS12] A. Mielke, C. Ortner, and Y. Sengül. Quasistativ nonlinear viscoelasticity as a curve of maximal slope. *In preparation*, 2012.

[MRS08] A. Mielke, T. Roubíček, and U. Stefanelli. Γ-limits and relaxations for rate-independent evolutionary problems. *Calc. Var. Part. Diff. Eqns.*, 31, 387–416, 2008.

[MuN08a] C. B. Muratov and M. Novaga. Front propagation in infinite cylinders. I. A variational approach. *Commun. Math. Sci.*, 6(4), 799–826, 2008.

[MuN08b] C. B. Muratov and M. Novaga. Front propagation in infinite cylinders. II. The sharp reaction zone limit. *Calc. Var. Partial Differential Equations*, 31(4), 521–547, 2008.

[Nay76a] B. Nayroles. Un théorème de minimum pour certains systèmes dissipatifs. Variante hilbertienne. *Travaux Sém. Anal. Convexe*, 6(Exp. 2), 22, 1976.

[Nay76b] B. Nayroles. Deux théorèmes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17), Aiv, A1035–A1038, 1976.

[Ons31] L. Onsager. Reciprocal relations in irreversible processes, I+II. *Physical Review*, 37, 405–426, 1931. (part II, 38:2265-227).

[Ott98] F. Otto. Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Rational Mech. Anal.*, 141(1), 63–103, 1998.

[Ott01] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26, 101–174, 2001.

[OtW05] F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Analysis*, 37, 1227–1255, 2005.

[PeR11] M. A. Peletier and M. Renger. Variational formulation of the Fokker-Planck equation with decay: a particle approach. *arXiv:1108.3181v1*, 2011.

[Pet04] J. Petersson. A note on quenching for parabolic equations with dynamic boundary conditions. *Nonlinear Anal.*, 58(3-4), 417–423, 2004.

[PoA06] J. Poortmans and V. Arkhipov. *Thin film solar cells: fabrication, characterization and applications.* Wiley series in materials for electronic and optoelectronic applications. Wiley, 2006.

[PuF97] S. Puri and H. L. Frisch. Surface-directed spinodal decomposition: modelling and numerical simulations. *Journal of Physics: Condensed Matter*, 9(10), 2109, 1997.

[RaZ01] R. Racke and S. Zheng. The Cahn-Hilliard equation with dynamic boundary conditions. *Advances Di. Equations*, 8, 8–83, 2001.

[Rou05] T. Roubíček. *Nonlinear partial differential equations with applications.* ISNM. International Series of Numerical Mathematics 153. Basel: Birkhäuser, 2005.

[RS*11a] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli. A variational principle for gradient flows in metric spaces. *C. R., Math., Acad. Sci. Paris*, 349(23-24), 1225–1228, 2011.

[RS*11b] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli. Weighted energy-dissipation functionals for gradient flows in metric spaces. 2011. In preparation.

[Sal84] A. Salvadori. On the M-convergence for integral functionals on L^p_X. *Atti Sem. Mat. Fis. Univ. Modena*, 33, 137–154, 1984.

[SaS04] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, LVII, 1627–1672, 2004.

[SaV97] G. Savaré and A. Visintín. Variational convergence of nonlinear diffusion equations: Applications to concentrated capacity problems with change of phase. *Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat., IX. Ser., Rend. Lincei, Mat. Appl.*, 8(1), 49–89, 1997.

[Sch94] D. Schroeder. *Modelling of interface carrier transport for device simulation.* Computational microelectronics. Springer-Verlag, 1994.

[ScT10] K. Schmidt and S. Tordeux. Asymptotic modelling of conductive thin sheets. *Z. Angew. Math. Phys.*, 61(4), 603–626, 2010.

[SeT12] E. Serra and P. Tilli. Nonlinear wave equations as limits of convex minimization problems: proof of a conjecture by De Giorgi. *Ann. of Math. (2)*, 175(3), 1551–1574, 2012.

[ShS98] J. Shatah and M. Struwe. *Geometric wave equations*, volume 2 of *Courant Lecture Notes in Mathematics.* New York University Courant Institute of Mathematical Sciences, New York, 1998.

[SpS11] E. Spadaro and U. Stefanelli. A variational view at the time-dependent minimal surface equation. *Journal of Evolution Equations*, 11, 793–809, 2011.

[SpW10] J. Sprekels and H. Wu. A note on parabolic equation with nonlinear dynamical boundary condition. *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods*, 72(6), 3028–3048, 2010.

[Ste08a] U. Stefanelli. The Brezis – Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.*, 47(3), 1615–1642, 2008.

[Ste08b] U. Stefanelli. A variational principle for hardening elastoplasticity. *SIAM J. Math. Anal.*, 40(2), 623–652, 2008.

[Ste09] U. Stefanelli. The discrete Brezis-Ekeland principle. *J. Convex Anal.*, 16(1), 71–87, 2009.

[Ste11] U. Stefanelli. The De Giorgi conjecture on elliptic regularization. *Math. Models Methods Appl. Sci.*, 21(6), 1377–1394, 2011.

[Str06] M. Struwe. On uniqueness and stability for supercritical nonlinear wave and Schrödinger equations. *Int. Math. Res. Not.*, pages Art. ID 76737, 14, 2006.

[Stu05] K.-T. Sturm. Convex functionals of probability measures and nonlinear diffusions on manifolds. *J. Math. Pures Appl. (9)*, 84(2), 149–168, 2005.

[SzN07] S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices.* Wiley-Interscience publication. Wiley-Interscience, 2007.

[Tai09] K. Taira. *Boundary value problems and Markov processes. 2nd ed.* Lecture Notes in Mathematics 1499. Berlin: Springer. xii, 186 p., 2009.

[Vil09] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009. Old and new.

[Vis96] A. Visintin. *Models of phase transitions.* Progress in Nonlinear Differential Equations and their Applications, 28. Birkhäuser Boston Inc., Boston, MA, 1996.

[Wł87] J. Włoka. *Partial differential equations. Transl. from the German by C. B. and M. J. Thomas.* Cambridge etc.: Cambridge University Press. XI, 1987.